Economic Decisions under Uncertainty

by Hans-Werner Sinn

North Holland: Amsterdam, New York and Oxford 1983

Chapter 3: The Structure of Risk Preference

Chapter Three
The Structure of Risk Preference

In the two preceding chapters it was shown how consistent decisions can be reached under uncertainty. By the use of, subjectively formed, equivalent objective probabilities the expected utility of each decision alternative has to be determined and then the alternative with the highest level of expected utility must be chosen. Unfortunately this advice remains quite meaningless as long as all that is known is that the utility function reflects the decision maker's preferences, while the form of this function is quite unknown. This chapter, therefore, attempts to gain more specific information on the shape of the utility function.

Part A sets out some scientific findings from psychophysics. Part B considers the evaluation of probability distributions involving negative wealth. Part C discusses a rival preference hypothesis formulated by Arrow.

For the following analysis, we need a more detailed idea of the notion of 'wealth' than before. Wealth is seen as an indicator of all future consumption possibilities, and it is assumed that these consumption possibilities are the source of the utility gained from owning wealth. Wealth, therefore, is defined as the sum of material wealth and human capital. Human capital is the present value of all future net incomes and is assumed to be non-random. In the classical manner 'net' means that the reproduction cost of labor, i.e., the physical subsistence minimum, is deducted from the income of labor. Accordingly, the future consumption possibilities mentioned above comprise gross consumption minus what is necessary for minimum subsistence. It is assumed that the path of consumption[1] being fed out of end-of-period wealth v is proportional to the size of this wealth. This assumption, which will be legitimated in

[1] In this chapter it is still assumed that, in the current period for which the decision under risk is examined, there is no consumption. This assumption is removed in chapter IV B.

the next chapter, is compatible with monetarist consumption hypotheses like those of MODIGLIANI/BRUMBERG (1955) and FRIEDMAN (1957). Moreover it is assumed that, at each point in time, the money value of consumption can be interpreted as a quantity index of a homogeneous bundle of consumption goods. This is a highly idealized assumption, since the structure of the bundle demanded will, in general, vary with its size. The idealization is roughly the kind used when national product is interpreted as a quantity measure[2] and may be justified by the fact that the evaluation by market prices transforms heterogeneous commodities into homogeneous values[3].

Section A
Psychological Aspects of Risk Evaluation

1. *Psychological Relativity Laws*

The next section will establish a hypothesis concerning the shape of the von Neumann-Morgenstern utility function. This section provides the key to this hypothesis by discussing theoretical and empirical aspects of a fundamental psychological relativity law.

1.1. *Bernoulli's Relativity Law*

Among the topics treated in BERNOULLI's (1738) *specimen theoriae novae de mensura sortis,* it is the expected-utility rule as such that particularly seems to interest contemporary economists, although this rule was actually developed by CRAMER (1728), not by Bernoulli himself. Bernoulli's own contribution was to establish a particular hypothesis regarding the shape of the utility function. This hypothesis is that the utility function is logarithmic.

---

[2] This interpretation is common practice in public discussions and underlies all macro economic one-sector models.

[3] Our assumption may be also explained by the theory of the *utility tree* developed by STROTZ (1957, 1959) and GORMAN (1959a and b). According to this theory, a decision maker first determines the optimal apportioning of his budget between different commodity bundles which are considered as homogeneous goods, the price of which is the price index of the single commodities in this bundle. Only after this decision are the subbudgets so formed apportioned. As suggested by the theory, end-of-period wealth can be perceived as representing a single homogeneous commodity bundle whose structure is not subject to choice at the present stage of decision making. Interpreted in this way, the intertemporal optimization approach presented in the next chapter deals with the decision problem on a higher level of the utility tree where wealth is no longer homogeneous but where consumption still is.

The reasoning in favor of this function is of such an amazing originality that it is surprising that so little attention has been paid to it in recent times, except for an appreciative remark by BORCH (1968a, p. 45) and SAVAGE's (1954, p. 94) statement 'To this day, no other function has been suggested as a better prototype for Everyman's utility function.' At the turn of the century, the situation was different. Then, the logarithmic function was seen as the crucial point in Bernoulli's article[1]. MARSHALL (1920, pp. 111f.) rather favored this function, though with some reservations; it was discussed in connection with the sacrifice theory of taxation by, for example, SAX (1892, pp. 76–79); and, still earlier, LAPLACE (1814, p. XV) had championed it.

Let us then follow Bernoulli's reasoning. After giving a definition of wealth largely compatible with the one given above[2], he says (§ 6):

> If a person '... has a fortune worth a hundred thousand ducats and another a fortune worth the same number of semi-ducats and if the former receives from it a yearly income of five thousand ducats while the latter obtains the same number of semi-ducats it is quite clear that to the former a ducat has exactly the same significance as a semi-ducat to the latter, and that, therefore, the gain of one ducat will have to the former no higher value than the gain of a semi-ducat to the latter. Accordingly, if each makes a gain of one ducat the latter receives twice as much utility from it, having been enriched by two semi-ducats.'

In these sentences Bernoulli formulates a psychological relativity law. For the rich man the gain of one ducat has the same meaning as the gain of half a ducat has for the man who is only half as rich since, from a subjective point of view, it is the *relative* change in wealth that matters.

The way Bernoulli's argument is formulated suggests that it is not limited to small wealth changes of one or two ducats but is meant in a broader sense. 'Equal percentage changes in wealth induce equal absolute changes in utility' seems to be what he wanted to say, i.e.,

$$\Delta U = f\left(\frac{\Delta v}{v}\right), \quad f'(.)>0, \quad f(0)=0. \tag{1}$$

The only reason for referring to small changes in wealth is to make plausible a differential equation of the type

$$dU = b\frac{dv}{v}, \quad b = \text{const.} > 0, \tag{2}$$

[1] See the comment made in 1896 by the editor of the German translation of BERNOULLI (1938), *L. Fick*, on p. 8 of his introduction.

[2] Even the problem of a homogeneous flow of consumption goods that is fed out of wealth is clearly seen by Bernoulli. Cf. his prisoner example in § 5.

for here, instead of the function symbol $f(.)$, a constant factor b can be used. Assume that the utility function $U(.)$, and hence also the function $f(.)$, is twice differentiable. Then, with $\Delta v/v \to 0$, (2) indeed follows from (1) if we set $b = f'(0)$. An integration of (2) immediately produces Bernoulli's function $U(v) = a + b \ln v$ or rather, because of the meaningless of a linear transformation[3],

$$U(v) = \ln v. \tag{3}$$

This in turn implies that the function $f(.)$ postulated in (1) does indeed exist and that it shows the desired properties for large changes in wealth as well as for small changes:

$$\begin{aligned} \Delta U = f\left(\frac{\Delta v}{v}\right) &= \ln(v + \Delta v) + \ln v \\ &= \ln\left(\frac{v + \Delta v}{v}\right). \end{aligned} \tag{4}$$

It is useful to calculate for Bernoulli's function the *intensity of insurance demand* as defined in chapter II C 1.3. Applying the inverse function $U^{-1}(.)$ to $U[S(aq - C)] = E[U(aq - C)]$, i.e., to $\ln[S(aq - C)] = E[\ln(aq - C)]$ we can derive an explicit expression for the certainty equivalent and thus for the intensity of insurance demand:

$$g = \frac{aq - \prod_{i=1}^{n} (aq - c_i)^{w_i}}{E(C)}. \tag{5}$$

Here c_i is the ith variate of the loss variable and w_i is its probability[4]. As can easily be seen, g is homogeneous of degree zero:

$$g\lambda^0 = \frac{\lambda aq - \prod_{i=1}^{n} (\lambda aq - \lambda c_i)^{w_i}}{E(\lambda C)}. \tag{6}$$

Thus Bernoulli's relativity law brings about the interesting result that, with given probabilities, a proportional increase in wealth and all loss variates also increases proportionally the maximum premium from the

[3] Cf. expression (II C 2).

[4] Because of (II C 14) and (II C 17), $g = [aq - S(aq - C)]/E(C)$, and from $\ln[S(aq - C)] = E[\ln(aq - C)]$ it follows that

$$S(aq - C) = \exp E[\ln(aq - C)] = \exp \sum_i w_i \ln(aq - c_i) = \prod_i (aq - c_i)^{w_i}.$$

point of view of the purchaser. In other words: the intensity of demand for wealth insurance is independent of the size of wealth.

This plausible implication of Bernoulli's relativity law was quite recently classified by PRATT (1964) and ARROW (1965) as *constant relative risk aversion* and would have been willingly accepted in modern risk theory had Bernoulli offered a more profound basis for his analysis. Quite correctly, as was shown in chapter II, he was critized for using the utility function of non-random wealth for the evaluation of probability distributions. But although the criticism only referred to this part of Bernoulli's argument, the argument as a whole was dismissed. In this way, quite unintentionally, the baby was thrown out with the bath water.

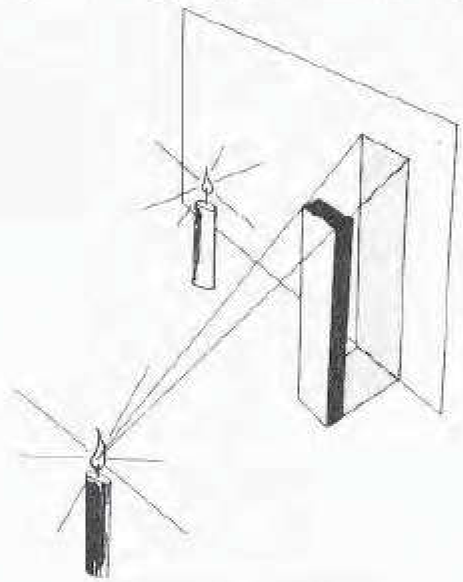
To show this it seems worth-while to digress to psychophysics where similar ideas have been developed. Psychophysics allows us to find the proper basis of Bernoulli's reasoning so that later we can attempt to rebuild his approach on a more solid foundation.

1.2. *The Relativity of Stimulus Thresholds*

If someone whispers softly we often cannot understand him even when everything is very quiet. When the surrounding noise level is high, we sometimes cannot hear what he says even if he speaks loudly. Even on a clear moonlit night our naked eye can only see a few of the stars we can see through a high-powered telescope. In daylight we can hardly see any stars at all although they are shining just as brightly as they are at night. In all these cases, stimulus thresholds obstruct proper perception. In the quiet room or at night absolute limits to sensation are encountered; thus the term *absolute threshold* is used. With high noise levels or in broad daylight the intensity of the stimulus is not sufficient to produce a perceptible difference in relation to the environmental stimulus; thus the term *differential threshold* is used.

The differential threshold is of particular interest. The above examples suggest the possibility that a given absolute change in the intensity of a stimulus is perceived when the overall intensity is low but not when it is high.

This phenomenon can be formulated more precisely by referring to an experiment carried out more than two hundred years ago by BOUGUER (1760, pp. 51–58). There is a dark room with a white screen. In front of the screen at a distance of one foot a lighted candle and a dark form are placed side by side. A second lighted candle is placed in front of the form so that the form is silhouetted against the screen. The distance between this candle and the screen is altered until it just becomes impossible to perceive the silhouette against the bright background. Bouguer found that there was a critical value of 8 feet, which means that the brightness of a shadow must be exceeded by the brightness of its environment by 1/64 if it is to be perceptible. FECHNER (1860 I, pp. 147–151) repeated the experiment and found a surprising invariance in the proportion 1/64 that Bouguer had indicated; whatever the distance between the screen and the first candle, the crucial distance for the second candle is always exactly eight times as large.



Bouguer's experiment

Figure 1

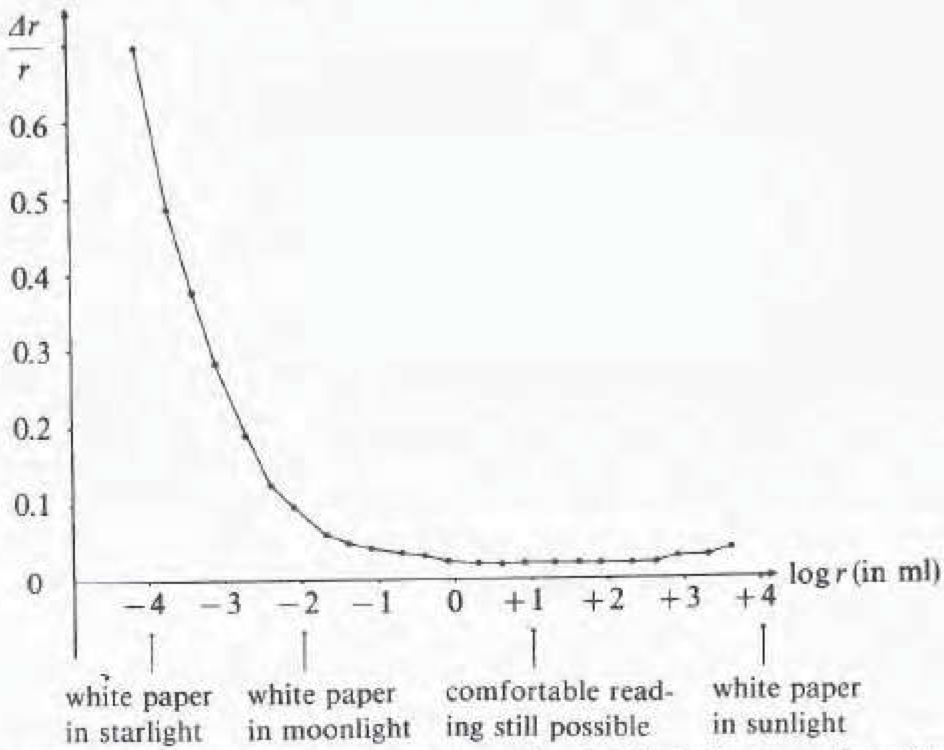
Let r denote the stimulus intensity and Δr the differential threshold. Then this result implies

(7) $$\Delta r = \alpha r, \quad \alpha = \text{const.},$$

or in words: *The differential threshold is a given proportion of the initial stimulus intensity.* This general phenomenon was called *Weber's law* by FECHNER (1860 I, pp. 134–139).

Independently of Bouguer, Ernst Heinrich WEBER (1834, esp. pp. 87–90, 171–175; 1846, esp. pp. 89, 104–107, 134–139) found similar results in a series of weight experiments he carried out. The question was how big must the difference between two weights be before they are perceived as being different when lifted manually. The remarkable result that Weber found is that the difference is a given proportion of the lighter or heavier weight, respectively, largely independent of the size of these weights. For the human hand Weber found a relative threshold of 1/40 with respect to the ligher weight. This result was confirmed in many further experiments. Weber also examined the human ability to distinguish between the lengths of two lines shown consecutively, and between two consecutively produced pitch frequencies. He found relative differential thresholds of 1/100 or 1/322, respectively, for persons experienced in the tests.

After Weber, many other scientists, FECHNER (1860 I and II)[5] in particular, experimented with a large number of stimulus continua[6], being all measurable on a ratio scale. The results of these experiments broadly confirm Weber's findings.



Source: KÖNIG and BRODHUN (1889) (results for König) in connection with information from HECHT (1934) and RIGGS (1971)

Figure 2

For a time it was, however, fashionable to question the validity of Weber's law. It was argued that the relative threshold is not quite so invariant with respect to the intensity of the stimulus as Weber's law suggested[7], for the curve relating the relative threshold to the stimulus intensity seems typically to have a shape similar to the one reported in Figure 2. This curve was constructed from data provided by KÖNIG and BRODHUN (1889). It reveals that the relative stimulus threshold for small

[5] Among the continua studied by Fechner are brightness, sound pressure, weight, visual estimation, tactile estimation, and color.

[6] For example KÖNIG and BRODHUN (1888, 1889; brightness); VON BÉKÉSY (1930; sound and vibration); HOLWAY and PRATT (1936; taste, loudness, smell, brightness). Recently Weber's law was shown to hold even for very small time intervals: GETTY (1975).

[7] See BORING (1942, pp. 138 f.).

levels of the stimulus intensity, i.e., for intensities close to the absolute threshold, is greater than for intensities in the medium range. The same holds for intensities close to the level which would destroy the receptor organ. For the light continuum, the deviation from Weber's law, which is shown by König and Brodhun's data, can be confirmed by anyone who thinks about the way contrasts seem to vanish when the light is either extremely bright or extremely dim. Thus Weber's law does not hold for extreme intensity levels. On the other hand, its practical validity is much greater than an initial glance at Figure 2 suggests. Note that the abscissa denotes the logarithm of stimulus intensity. As soon as we plot the same figure with reference to the intensity itself the range where the relative threshold is constant immediately appears to be enormously expanded. This appearance is not deceptive for, as STEVENS (1951, p. 35) estimates for the cases of sound and light sensation, the horizontal range covers 99.9% (!) of the intensities that occur in practice.

Of course, there are additional problems. The stimulus threshold will generally depend on a number of other factors such as fatigue, practice, mood, etc. If these influences cannot be taken into account, they lead to stochastic disturbances in Weber's law[8]. This, however, does not affect the general phenomenon that, under given exogenous conditions, there is a given relative threshold that, over the practically relevant range, is independent of stimulus intensity.

1.3. *The Psychophysical Law*

Gustav Theodor Fechner devoted his life's work to the discovery of the relationship between body and soul. Fortunately he reduced this highly metaphysical intention to the problem of finding a functional relationship between the intensity of a stimulus and the intensity of sensation, i.e., to finding the so-called *psychophysical law*. Thus, Fechner became the founder of a branch of science, which, at the end of the last century, was intensively investigated in Germany and which has had a renaissance in recent times through the experimental work of Stanley S. Stevens and his co-workers. FECHNER's (1860 I and II) *Elemente der Psychophysik* and STEVENS's (1975) *Psychophysics* are landmarks in scientific research set a century apart.

1.3.1. Fechner's Law

Searching for the psychophysical law

(8) $$s = s(r),$$

[8] Cf. the foundation of stochastic scaling methods by THURSTONE (1927).

where s denotes the subjective intensity of a physical stimulus of intensity r, Fechner made use of Weber's threshold theory that he had himself confirmed by a number of experiments which are still famous today. He assumed that every just perceptible change in a stimulus intensity brings about the same change in the intensity of sensation. Accordingly, he argued that the difference between two levels of sensation intensity can be measured by the number of just perceptible steps between them. Let α denote the ratio between the differential threshold and the corresponding previous level of stimulus intensity. Then, for two stimulus intensities r and r^*, $r > r^*$, between which there are n steps[9], it holds that

$$(9) \qquad r = (1+\alpha)^n r^*.$$

Here, by assumption,

$$(10) \qquad n \equiv s(r) - s(r^*)$$

denotes the difference in sensation intensities which is defined up to a multiplication with a positive constant. Inserting (10) into (9) and taking the logarithm we have

$$(11) \qquad s(r) - s(r^*) = \ln\left(\frac{r}{r^*}\right)\frac{1}{\ln(1+\alpha)}.$$

This expression implies a result that is denoted as Fechner's law: *Equal relative changes in stimulus intensity bring about equal absolute changes in sensation intensity*. With $r^* = \text{const.}$ (11) gives

$$(12) \qquad s(r) = \alpha + \beta \ln r, \ \beta > 0,$$

an equation describing the relationship between stimulus and sensation we were looking for[10].

[9] Cf. FECHNER (1860 II, pp. 9–29). We do not follow Fechner's presentation. The integration of his fundamental equation $dG = b\,dr/r$ which corresponds closely to that of Bernoulli (cf. equ. (2)) must be interpreted as an approximation since thresholds cannot become infinitesimally small. This approximation is, in principle, admissible but at this stage it is avoidable.

[10] Note that, because of the way it was derived, $s(r)$ is only defined for a discrete scale with given sensation intervals as units. Thus it seems that our senses show the environment in a digital rather than in an analogous way. This observation, seen in the light of what we know about digital computers, suggests a possible explanation of thresholds as such. Thresholds reduce precision in monitoring the environment. Such a reduction in precision was an evolutionary advantage which developed in order to economize on the 'costly' calculation capacities of our brains. Cf. the discussion of the Axiom of Ordering in connection with thresholds (ch. I 1.1).

Equation (12) and the reasoning behind it are very similar to Bernoulli's theory. FECHNER (1860 I, esp. pp. 236–238) knew this. He went so far as to stress the appropriateness of his formula for a utility evaluation of wealth, arguing that wealth is a means of producing 'werthvolle Empfindungen' (valuable sensations) and so can act like any other stimulus.

Fechner's law was subject to a great deal of criticism, for he did not attempt to legitimate his basic assumption that all just perceptible increases in stimulus intensity appear subjectively equal. So his law does *not* follow automatically from Weber's law and it is easy to understand why, for a long time, it was not accepted.

1.3.2. Stevens's Law

Stevens was one of Fechner's most profound critics. After the most extensive experimental work in the field of scaling problems that has ever been carried out, he believed he had good reason to replace Fechner's logarithmic law by a *power law*[11]. The sensation function that he found inductively is determined up to a multiplication with a strictly positive factor of proportionality[12] κ and reads

$$(13) \qquad s(r) = \kappa r^{\Theta};\ \kappa, \Theta > 0.$$

Because of

$$(14) \qquad \ln s = \ln \kappa + \Theta \ln r$$

it is a straight line in a $(\ln s, \ln r)$ diagram. The power function describes another psychophysical relativity law. Since, comparing two stimulus intensities r and r^*, we have

$$(15) \qquad \frac{s(r)}{s(r^*)} = \left(\frac{r}{r^*}\right)^{\Theta},$$

Stevens's law can be formulated analogously to Fechner's law[13]: *Equal*

---

[11] Cf. in particular STEVENS (1956, 1959, 1961, 1962, 1966, 1975). The survey of his works published in 1975 was finished by Stevens shortly before his death.

[12] The proportionality factor has no meaning since a change in the unit in which r is measured changes its size:

$$\kappa'(rx)^{\Theta} = \kappa r^{\Theta} \quad \text{with} \quad \kappa \equiv \kappa' x^{\Theta}.$$

[13] STEVENS (1975, p. 16). This formulation had already been used in the discussion of Fechner's law that took place at the end of the last century. Cf. FECHNER (1888, pp. 174–179).

relative changes in stimulus intensity bring about equal relative changes in sensation intensity.

Stevens's law had previously been derived by PLATEAU (1872), FULLERTON and CATTEL (1892)[14], MEINONG (1896), and GUILFORD (1932) from a modification of Weber's law[15]. However the legitimation provided by these authors met with little approval.

Thus Stevens wanted his sensation function to be seen independently of Weber's law which he regarded as incontestable. Accordingly, his experiments were not designed to find out about thresholds in order to measure sensation indirectly via the summation of these thresholds. Instead, the persons interviewed were asked to indicate the magnitude of alternately offered stimulus intensities by numbers they themselves could choose. This direct method of measuring sensation is called *number matching*.

Table 1
Estimates in the Size of Selected Stimuli*
(measured exponent in $s = \kappa r^{\Theta}$)

loudness (sound pressure of 3000 Hz)	0.67
vibration 250 Hz	0.6
vibration 60 Hz	0.95
brightness (point source)	0.5
visual length	1.0
visual area	0.7
saturation of a red-gray mixture of colors	1.7
salt concentration	1.4
muscle force	1.7
heaviness	1.45
electric shock	3.5
vocal effort	1.1

**Source:* STEVENS (1975, p. 15).

Those who, along with Pareto, Hicks, and Allen, believe that people cannot do more than indicate 'better or worse', 'larger or smaller', and 'more or less' will be surprised to hear about Stevens's findings. People did not feel overburdened when asked to assign numbers to the stimuli offered to them: they acted in the way described by Stevens's law, and not just by making stabs in the dark. Table 1 reports some of the exponents that Stevens measured with the method of number matching.

The reliability of Stevens's results can be confirmed by a modification

[14] Cited according to GUILFORD (1932, pp. 40 f.).

[15] FECHNER (1877, pp. 10 ff. and 21) characterized the difference between the logarithm and the power laws by distinguishing the equations $ds = \beta\, dr/r$ and $ds/s = \Theta\, dr/r$. A comparative discussion of the two approaches is given by GROTENFELT (1888, pp. 15–20), WUNDT (1908, pp. 638–645), and MEINONG (1896, pp. 380–388).

of the experiments presented by STEVENS (1959, 1961, and 1966) that is called *cross-modality matching*[16]. According to this method, people are asked to associate different intensities of a stimulus not with numbers but with the intensities of another, freely choosable, stimulus in such a way that both stimuli seem equally intensive. Thus, with the two sensation functions $s_1 = \kappa_1 r_1^{\Theta_1}$ and $s_2 = \kappa_2 r_2^{\Theta_2}$, for any r_2 offered, r_1 must be chosen by the experimental subject so as to ensure

$$(16) \qquad s_1 = s_2 .$$

Because of

$$(17) \qquad \kappa_1 r_1^{\Theta_1} = \kappa_2 r_2^{\Theta_2}$$

and hence

$$(18) \qquad \ln r_1 = \frac{\ln \kappa_2 - \ln \kappa_1}{\Theta_1} + \frac{\Theta_2}{\Theta_1} \ln r_2$$

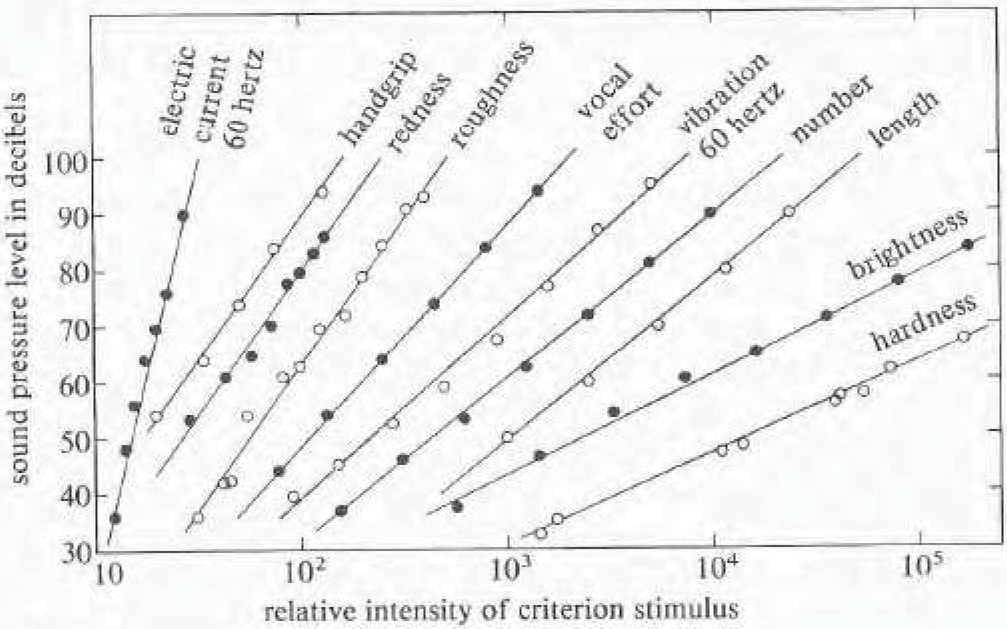
a straight line has to be found in a $(\ln r_1, \ln r_2)$ diagram if Stevens's law is to be valid. In addition, the slope of this line has to equal the ratio Θ_2/Θ_1 of the slopes found by the method of number matching. To a surprising extent, both requirements are met[17]. An example for the result of a cross-modality study is illustrated in Figure 3. There, sound pressure acts as stimulus r_1 and the other stimuli indicated act alternately as r_2. Because of the irrelevance of their vertical intercepts, the lines in Figure 3 are arbitrarily ordered, but the slopes reflect the measurement results. Since the number-matching exponents are known for all stimulus intensities r_2, it is possible to calculate estimates $\hat{\Theta}_1 = \Theta_2/\Theta^*$ for the loudness exponent with the aid of the slopes Θ^* of the curves plotted in Figure 3. Stevens found a geometric mean of 0.67. This value is almost identical to the exponent measured in direct number-matching experiments[18]. Perhaps chance was responsible for this surprising result. However, STEVENS (1975, pp. 113 and 117) succeeded in citing a number of further cross-modality experiments that led to a fairly consistent structure of exponent ratios. By the standards of a social science theory this result is of exemplary accuracy.

---

[16] Cf. STEVENS (1975, pp. 109–111). The method seems to have first been used by VON BÉKÉSY (1930, pp. 346–348) who compared the sensations of vibration and loudness. VON BÉKÉSY found a value of $\Theta_1/\Theta_2 = 1$ which roughly fits the results reported in Table 1.

[17] See STEVENS (1975, pp. 99–133).

[18] STEVENS (1975, p. 119) cites a study by Moskowitz (1968), not available to the author, where a value of 0.676 is found.



Source: STEVENS (1966).

Figure 3

1.3.3. The Missing Numeraire

Although there is no reason to doubt the validity of Stevens's empirical findings[19], different interpretations are possible. Stevens himself proposed one, namely, that the subjective intensity of sensation resulting from an objective intensity of a stimulus was measured. But there is another possibility[20]. It traces back to a short note by EKMAN (1964) and is based on previous conjectures by GARNER, HAKE, and ERIKSEN (1956, pp. 155–157) and ATTNEAVE (1962, pp. 623–627). This interpretation is that number-matching is a special kind of cross-modality matching[21], since the persons participating in the number-matching experiment set the intensity of number sensation equal to the intensity of the stimulus proper. Suppose in equations (16)–(18) the function $s_1 = \kappa_1 r_1^{\Theta_1}$ is the sensation function for numbers and suppose further that, through number matching, the exponent Θ^* has been found for a certain stimulus r_2. Then this exponent is in fact the ratio of

---

[19] The author is not aware of any criticism of the reliability of Steven's experiments.

[20] A further interpretation was given by WARREN (1958) who argued that the measured result reflects the correlation between the offered stimuli as experienced in reality and/or knowledge of physical scales if available. This interpretation is not compatible with the fact that experimental subjects make 'mistakes' when estimating well-known continua such as areas and weights (cf. Table 1). Moreover, this interpretation cannot explain the fact that cross-modality matching brings about consistent results even when people are required to match stimuli with one another that, in real life, are not correlated.

[21] STEVENS (1975, pp. 34 and 107 f.) is sympathetic to this interpretation without, however, drawing conclusions similar to those of Ekman.

the true exponent of the sensation function for the stimulus in question (Θ_2) and for numbers (Θ_1):

$$(19) \qquad \Theta^* = \frac{\Theta_2}{\Theta_1}.$$

Since of course a similar result holds for all other estimated exponents, in all cases one scale is merely measured by the other. The whole system of cross-modality estimations lacks a numeraire that provides the link with the true intensity of sensation.

The only, by no means convincing, 'argument' Stevens was able to put up against this view was simply the assumption that the exponent of number sensation is unity[22]. This assumption does not become any more plausible merely because the number continuum and the length of a straight line are subjectively proportional ($\Theta^* = 1$)[23], as Table 1 reveals. For how do we know what the law is that governs the sensation intensity when the length of a straight line is perceived? From Fechner's point of view, it could be argued that the application of Weber's law to a comparison of distances establishes a, strongly curved, logarithmic sensation function and not a linear one.

How much the lack of an anchor causes the position of Stevens's system of exponents to drift may be shown by a thought experiment. Suppose the true exponent of number sensation falls from 1 to 0 so that, given the exponent ratios, all number-matching exponents also fall to zero. In a ($\ln s, \ln r$) diagram like that of Figure 3, this rotates all the straight lines towards the abscissa. Given the range of values on the $\ln r$ axis the range of values on the $\ln s$ axis then shrinks to zero. Thus the curvature of the $\ln s$ curve progressively loses its significance and the curve may finally be approximated by a straight line. Thus, in practical terms, we approach a semi-logarithmic diagram which implies that, in the limiting case of our thought experiment, *all* of Stevens's power functions reduce to logarithmic functions of the Fechner type.

To check this result algebraically assume that the true sensation functions are logarithmic and try $s_1 = \alpha_1 + \beta_1 \ln r_1$ for the number sensation function and $s_2 = \alpha_2 + \beta_2 \ln r_2$ for any of the other functions. Because of[24]

$$(20) \qquad \alpha_1 + \beta_1 \ln r_1 = \alpha_2 + \beta_2 \ln r_2$$

[22] STEVENS (1975, p. 107).

[23] STEVENS (1975, p. 14).

[24] Here the logarithmic functions are defined up to an additive constant since a change of the dimension of r must not exhibit any influence:

$$\alpha' + \beta \ln(rx) = \alpha + \beta \ln r \quad \text{with} \quad \alpha \equiv \alpha' + \beta \ln x.$$

the linear equation

$$\ln r_1 = \ln \kappa + \Theta \ln r_2 \quad \text{with} \quad \kappa = \exp\left(\frac{\alpha_2 - \alpha_1}{\beta_1}\right) \quad \text{and} \quad \Theta = \frac{\beta_2}{\beta_1}, \tag{21}$$

is indeed achieved. The equation corresponds to Stevens's function (14) if the numbers nominated by the people participating in the experiments are interpreted as information about stimulus r and not about sensation s. This is Ekman's result that, in a completely analogous way, can be extended to other combinations of stimuli.

A possible question arising at this stage is whether Stevens may be partially right since some functions are of the power type and others are logarithmic. The above heuristic thought experiment clearly answers this question in the negative. Indeed, trying $s_1 = \alpha_1 + \beta_1 \ln r_1$ and $s_2 = \kappa_2 r_2^{\Theta_2}$ we find that cross-modality matching requires

$$\alpha_1 + \beta_1 \ln r_1 = \kappa_2 r_2^{\Theta_2} \tag{22}$$

or equivalently

$$\ln r_1 = -\frac{\alpha_1}{\beta_1} + \frac{\kappa_2}{\beta_1} r_2^{\Theta_2}. \tag{23}$$

Since $r_2^{\Theta_2} = (e^{\ln r_2})^{\Theta_2} = e^{\Theta_2 \ln r_2}$ and $\Theta_2 > 0$ this implies a strictly convex curve in the $(\ln r_1, \ln r_2)$ diagram which is incompatible with Stevens's results. Thus Stevens's interpretation of the body of empirical findings must be either right or wrong. An intermediate solution does not exist. Summarizing, we may therefore state that whether the sensation functions belong to the class of power functions ($s = \kappa r^{\Theta}$) or to the class of logarithmic functions ($s = \alpha + \beta \ln r$) remains an open question despite the careful empirical research that has been carried out. In any case, all the functions must belong to the same class. For example, this means that all sensation functions are logarithmic if even a *single* one can be shown to be of this type.

1.3.4. Fechner's Law versus Stevens's Law: The Empirical Evidence

The Phenomenon of Logarithmic Interval Scales

The question we consider now is which class of sensation functions prevails in reality: the logarithmc or the power class. Neither adding just noticeable differences nor direct number matching provides an answer to this question. There is, however, a method that, at least in principle, can determine the correct sensation function.

This is the method of *interval* or *category estimation*. Here, the experimental subject is asked to classify given stimulus intensities into equidistant magnitude *categories* or to manipulate a set of stimulus intensities so that the distances between them seem to be subjectively equal. That this procedure cannot bring about more than an interval scale is self-evident. The basic difference between interval estimation and the direct methods of measuring employed by Stevens is that, rather than comparing two different continua, the increase in stimulus intensity on a certain level is compared with an increase in the intensity of the *same* stimulus on another level.

To facilitate an interpretation of the empirical findings, it is useful to consider the relationship between the subjectively equal distance Δr and the stimulus intensity r that would prevail under the two laws. As a first approximation we have

(24) $$\Delta s \approx s'(r)\Delta r = \text{const.}$$

Let $\eta_{\Delta r, r}$ denote the elasticity of Δr with respect to r and $\eta_{s'(r), r}$ the elasticity of $s'(r)$ with respect to r. Then, (24) implies that

(25) $$\eta_{\Delta r, r} = -\eta_{s'(r), r}.$$

The negative elasticity $-\eta_{s'(r), r}$ is a measure of concavity and uniquely characterizes the class of functions prevailing:

(26) $$-\eta_{s'(r), r} = \begin{cases} 1-\Theta \Leftrightarrow s(r) = \alpha + \beta\Theta r^{\Theta}, & \beta > 0, \quad \Theta \neq 0, \\ \quad 1 \quad \Leftrightarrow s(r) = \alpha + \beta \ln r, & \beta > 0. \end{cases}$$

With a little calculation, this contention can easily be checked. Combining (26) and (25) we find that the two competing laws imply

(27) $$\eta_{\Delta r, r} \begin{cases} <1, & \text{if} \quad s = \kappa r^{\Theta}, \Theta > 0 \quad \text{(Stevens)}, \\ =1, & \text{if} \quad s = \alpha + \beta \ln r \quad \text{(Fechner)}. \end{cases}$$

Thus, under *Fechner's law* the size of the subjectively equal increases in stimulus intensity rises *proportionally* and under *Stevens's law less than proportionally* with the objective level of stimulus intensity. If $\Theta < 0$, i.e., $\eta_{\Delta r, r} > 1$, then the sensation function is even more curved than the logarithmic function: neither Fechner's law nor Stevens's prevails.

Ekman's result showed that, from a theoretical point of view, there is no conflict between Stevens's empirical findings and Fechner's law. The rehabilitation of Fechner's law thus begun is completed by the empirical results achieved by using interval measurement. These assign far more relevance to the logarithmic sensation law than Stevens and his followers[25] have been willing to accept.

The first interval experiment was carried out by PLATEAU (1872). Asking painters to blend colors so as to produce a gray halfway between black and white he found a value of[26] $\eta_{\Delta r, r} = 2/3$. However, he was soon corrected by DELBOEF (1873, esp. pp. 50–101) who repeated the experiment in a somewhat modified form. Delboef produced the gray by means of a rotating disk with black and white areas, a procedure that has the advantage of giving more precise information about the proportions of the blend. His results favored the logarithmic function.

Repeating Delboef's experiment GUILFORD (1936, pp. 199 f.) found that the curvature of the sensation function is not, as Plateau contended, smaller than that of the logarithmic function, but is, on the contrary, bigger. From the numerical results he reports, the value of $\eta_{\Delta r, r} = 1.15$ can be calculated[27]. This again is closer to Fechner's than to Stevens's hypothesis.

HELSON's (1947) experiments also confirm the logarithmic function for the sensations of brightness and loudness. If a number of stimulus intensities are sequentially offered to the experimental subject then, in general, the geometric mean of the perceived intensities serves as the *adaptation level,* i.e., as the point of reference for subsequently offered intensities. The formula for the adaptation level (AL) of perceived stimulus intensities r_i is

$$\text{(28)} \qquad \mathrm{AL} = \prod_{i=1}^{n} r_i^{w_i}$$

where w_i denotes the weight factor of a particular stimulus. Because of

$$\text{(29)} \qquad \ln \mathrm{AL} = \sum_{i=1}^{n} w_i \ln r_i$$

---

[25] Cf. the articles contained in the 'Handbook of Perception' edited by CARTERETTE and FRIEDMAN (1974).

[26] If there are only two intervals that have to be set equal, the term *bisection method* is used. The experimental subject is asked to manipulate the intensity $\bar{r}$ of a stimulus so that it seems to be in the middle of two intensities offered to him. If Fechner's law is valid then it has to be expected that $\bar{r} = \sqrt{r_1 r_2}$, for this equation implies that

$$\ln \bar{r} = \frac{\ln r_1 + \ln r_2}{2}.$$

[27] Let the sensation function that is defined up to a positive linear transformation be $s = \Theta r^{\Theta}$, $\Theta \neq 0$, where the factor Θ merely has the task of determining the sign. Then for two stimulus intensities r_1 and r_2, whose psychological mean is $\bar{r}$, we have the formula $\Theta \bar{r}^{\Theta} = (\Theta r_1^{\Theta} + \Theta r_2^{\Theta})/2$ from which Θ and hence $-\eta_{s'(r), r} = \eta_{\Delta r, r} = 1 - \Theta$ can be calculated by a process of trial and error. For the estimates $r_1 = 100$, $r_2 = 2500$, and $\bar{r} = 411$ reported by Guilford a value of $\Theta = -0.1529\ldots$ is found.

this formula obviously implies a logarithmic sensation function[28].

A logarithmic function ($\eta_{\Delta r, r} = 1$), moreover, is the result of an experiment where the experimental subject has the task of ordering weights into equidistant categories. This experiment was reported by TITCHNER (1905a, pp. 33 f. and pp. 82–85) who dated it back to *Sanford*.

An *experimentum crucis* is a frequency test carried out by THURSTONE (1929) and GUILFORD (1954, pp. 103–106) that could have been constructed as an answer to Stevens's number matching except that the sequence in time makes this impossible. The experimental subjects are required to sort white cards, covered with black dots in different densities, into subjectively equidistant categories that are numbered consecutively. The result is that the category number is a logarithmic function of the true number of dots on the cards. An objection to this experiment, which equally well can be raised against Sanford, may be made[29] on the grounds that people possibly tend to fill the categories equally, so that the distribution of dot densities in the set of cards is crucial for the result of the experiment. However, even if this objection against logarithmic sensations of numbers were substantial, the question remains of why people chose a number system where the length of the written number is proportional to its logarithm rather than to the frequency it describes.

Despite all their criticism of the experiments of Thurstone and Guilford, even GALANTER (1957) and STEVENS (1961) confirm the tendency of these results[30]. Comparing number-matching scales and interval scales, they too find that the latter are biased towards a stronger curvature which, because of (24)–(27), is a bias towards Fechner's law.

After the studies of Galanter and Stevens, a number of further investigations into the size of this bias have been carried out. In their review article EKMAN and SJÖBERG (1965, p. 464) summarize the result that is a triumph for Fechner: 'The logarithmic relation between indirect interval and direct ratio scales is now a well-established fact for a great number of continua.' Among the empirical investigations, those of GALANTER and MESSICK (1961) and EISLER (1962b) in particular are worth stressing[31]. Both investigations show that for loudness the interval scale is logarithmic and the number-matching scale is of the power type. The result is of considerable significance in that, with respect to loudness in

---

[28] A detailed foundation for this function is given by HELSON (1964, pp. 57–62). The relationship to Fechner's law is treated in the 'reformulation' on pp. 197–231. Cf. also JOHNSON (1955, pp. 343–348).

[29] See THURSTONE (1929, pp. 223 f.).

[30] Cf. also STEVENS (1975, pp. 130 and 147–149).

[31] See STEVENS (1975, pp. 115–120).

particular, a great many cross-modality experiments have been carried out[32]. Recall however, that all scales in Stevens's system must be logarithmic if even one of these is shown to be logarithmic!

Apart from the results reported above, there is some evidence for logarithmic functions in connection with technical scales. For example, the decibel scale for loudness represents a logarithmic relation with physical sound pressure and the DIN scale of film speed has a corresponding property[33].

It is even more surprising that there is a logarithmic trend relationship between the musical scale and sound frequency as shown by Figure 4[34]. The reason for the systematic oscillation around the trend is that, although each successive octave doubles the frequency, the intervals between successive notes in the octave do not bring about the same percentage increases in frequency. The oscillation represents the *Wohltemperatur* (equal temperament) of the scale established by *J.S. Bach* which ensures that, to avoid fluctuations in pitch, the ratio of any pair of notes is an integer. Equal temperament does not contradict the Fechner hypothesis that equal relative changes in frequency seem to be

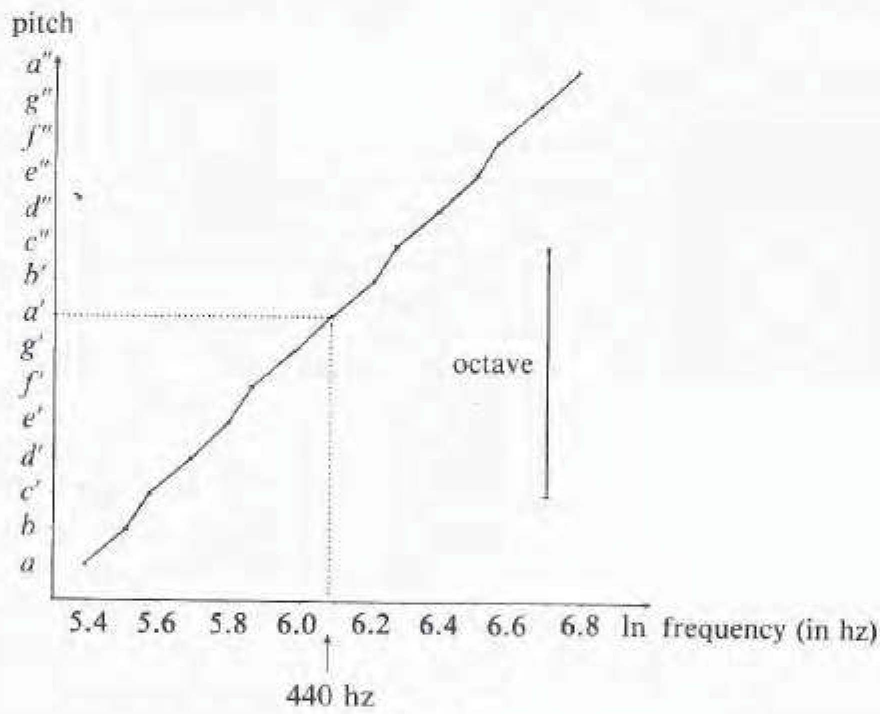


Figure 4

[32] Further experiments leading to analogous results were reported in Ekman and Sjöberg (1965).

[33] 3° DIN ≙ doubling of light intensity.

[34] For the historical development of the musical scale see Boring (1942, pp. 312–332). The connections with the psychophysical relativity law were already stressed by E.H. Weber (1846, p. 106), Wundt (1863, p. 81), and Lipps (1905, pp. 115–128).

subjectively equal, for the intervals on an equally tempered scale in fact appear to be different. This is demonstrated by the fact that the sound pattern of a melody played in A Major differs from that of the same melody played in C Major while a jump of an octave leaves the sound pattern unchanged.

The result is of particular importance in that mathematicians do not seem to have been involved in the historical development of the musical scale. Fechner's law is demonstrated very clearly: equal relative changes in sound frequencies are perceived as equal absolute changes in pitch.

This has also been confirmed by an experiment carried out by WARD (1954, esp. p. 373). He found that musicians have some difficulty in producing the octave of a basic note offered to them, but, when asked to produce even higher octaves, they stick to their subjective octave with a remarkable degree of accuracy. The subjective octave is in general not a change in frequency of 100% but is a *given* percentage[35] of the initial frequency.

The Results of Neurological Measurement

The empirical findings reported so far have one thing in common. In all cases they refer to a functional relationship between an objectively measurable quantity and a subjective intensity of sensation consciously revealed by experimental subjects. Another way of finding out about the intensity of sensation is to measure directly the electro-chemical processes in the nervous system. For the case of simple stimuli that, in principle, can be transmitted by single receptor organs, there is a substantial body of empirical evidence provided by the studies of FRÖHLICH (1921), ADRIAN (1928)[36], and many subsequent authors. It is now known that the stimulus arriving at a receptor is first transformed into an action current which is then transmitted via the nerve fibers in a complicated process of electrical and chemical reactions. What we should know about this process in order to evaluate the empirical results is that the intensity of a stimulus affects the intensity of the action current (as measured in volts) and the action current controls the frequency of impulses transmitted via the nerve fibers.

---

[35] Pitch belongs to the group of metathetic (qualitative) continua for which, according to STEVENS (1957), even number matching produces a logarithmic function. EISLER (1963, p. 252) remarks that this aspect implies a linear function for number sensation which is not compatible with the above statement that the function is logarithmic. EKMAN and SJÖBERG (1965, p. 470) object to this argument because metathetic scales do not have a subjective origin. If the experimental subject is nevertheless required to match numbers, they maintain, the numbers are used not for estimating magnitudes but for labelling categories of equal size in order to do 'the best in an impossible experiment situation'.

[36] For a summarizing discussion see also ADRIAN (1932, 1947).

Of particular interest are the functional relationships between the intensity of the stimulus, the intensity of the action current, and the impulse frequency. The first result cocerning these aspects was derived by FRÖHLICH (1921, esp. p. 15). He found a logarithmic relationship between the intensity of light and the intensity of the action current. Similarly, it was shown by HARTLINE and GRAHAM (1932), HARTLINE (1938), FUORTES (1959), and FUORTES and POGGIO (1963) that the way light intensity is transformed into impulse frequency is described by a logarithmic function. The compatibility with Fröhlich's result then obviously requires that the intensity of the action current and the impulse frequency vary in strict proportion. Precisely this was shown by KATZ (1950) and FUORTES and POGGIO (1963). These results could also be confirmed for other kinds of stimuli. For example, GALAMBOS and DAVIS (1943) and TASAKI (1954) found that loudness is approximately transformed into impulse frequency according to a logarithmic function and, according to MATTHEWS (1931) and VAN LEEUVEN (1949), the impulse frequency in those nerves that signal muscle tension is a logarithmic function of the weight carried by the muscle[37]. The parallelism between these results and Fechner's law, which has also been stressed by GRANIT (1955, pp. 8–23), cannot be overlooked[38].

---

37 Experiments with persistent stimuli show that, with the passage of time, the impulse frequency declines which is a sign of an adaptation process. The logarithmic functions usually refer to the maximum frequency defined as '1/(minimum time elapsing between two impulses)' or 'number of impulses in the first 1/10 second'. If the impulse frequency is measured over a longer period or after the passage of a given period of time then, in some experiments (Hartline), the logarithmic function no longer shows up since, in this case, a particular dependency between the speed of the adaptation process and the stimulus intensity affects the results. A similar problem arises when the frequency is measured for 'the first x impulses' since in this case there is a change even in the measurement period induced by a change in stimulus intensity. If the frequency is measured after the adaptation process, then it is again a logarithmic function of the stimulus intensity. Cf. GALAMBOS and DAVIS (1943, p. 48). It should be noted that one of the results reported by these authors (p. 47, Figure 8) implies a somewhat concave curve in a semi-logarithmic diagram and hence indicates a frequency function that is not only more curved than Stevens's power function but also more curved than Fechner's logarithmic function.

38 ROSNER and GOFF (1967) contend that the results are also compatible with the power law. Their own measurements, however, hardly support this view. They measure the relationship between the intensity of electric current (r) perceived by the experimental subject and the intensity of the induced electric current (s) in the brain, and plot their results in a diagram with the axes $\ln s$ (ordinate) and $\ln r$ (abscissa). Since all clouds of dots derived in this way very clearly suggest concave curves in this diagram it is to be expected that the authors do not find a power function but possibly a logarithmic function. They check both and indeed, for the latter, they calculate a lower variance of residuals (p. 201). They find, however, the smallest variance of residuals for a curve that is composed from two linear segments. If two linear segments had not brought a lower variance than the logarithmic function, they could also have chosen three or more. At some stage in this procedure they definitely would have reached a restatement of Stevens's law. What an excellent method!

It must be conceded, though, that in the studies cited above the true shape of the curve relating impulse frequency to stimulus intensity deviates from the logarithmic function in that, near the limits of the stimulus continuum, it is flatter than elsewhere. The shape resembles a curve that could be constructed by integrating the curve of König and Brodhun illustrated in Figure 2. Thus it seems that in this case the same phenomenon shows up that is found in threshold experiments and that was seen to be unimportant for the practically relevant range of stimulus intensities[39].

1.3.5. Result

The question about the *psychophysical law* is the question about the relationship between the objective intensity of a stimulus and the subjective intensity of its sensation. There are two answers competing with one another, *Fechner's logarithmic law* and *Stevens's power law*.

Fechner's law follows from Weber's law of a relativity in thresholds, when it is assumed that changes in stimulus intensity that just exceed a threshold are subjectively equal. The missing foundation of this assumption is the weakness of Fechner's law hypothesis. In contrast to this indirect way of reasoning, Stevens's law follows from number-matching experiments where people are asked to directly assign numbers to stimulus intensities offered to them. A consistent structure of sensation functions for a large number of stimulus continua has been built up as a result of number-matching experiments. Included are stimuli in a very broad sense, such as the length of a line and the size of an area. However, Stevens's law suffers from a serious drawback: it has to be assumed that the numbers chosen by the experimental subjects do, in fact, measure subjective sensation. If there is a subjective sensation function for numbers then all measures are cross-modality results so that a numeraire connecting Stevens's structure of power functions to true sensations is missing. Thus, a variety of different shapes for the unobserved true sensation function is compatible with Stevens's empirical findings. Among the possibilities are power functions, just as Stevens contended, but logarithmic functions are also possible. Only one thing can be firmly established: if Stevens's measurements are reliable, all functions must belong to the same class, i.e., for example, all functions are logarithmic if even one of them can be shown to have this property.

---

[39] That at its ends the empirical curve is flatter than the logarithmic curve is to be expected for purely technical reasons since there is an absolute lower threshold and an upper limit for the impulse frequency. The latter results from the fact that, after transmitting an impulse, a nerve cell has a phase of some 0.001 seconds during which it is unable to transmit a further impulse.

The theoretically appropriate way of measuring sensation is to use the method of interval estimation where the experimental subject is asked to order given stimulus intensities into equidistant categories or to produce increases in stimulus intensity that appear to be subjectively equal. The experiments carried out in this way confirm the hypothesis of logarithmic sensation functions, i.e., Fechner's law. If the results of these experiments as well as those achieved by Stevens are accepted, then only one conclusion is possible: even for all continua examined by Stevens there are logarithmic sensation functions and, in particular, there is a *logarithmic sensation function for numbers.*

In addition to the results from interval estimation there is further evidence in support of Fechner's law. This evidence is provided in neurological measurements of the relationship between stimulus intensity and the frequency of electrical impulses in nerve fibers. The results of these measurements are that impulse frequency is a logarithmic function of stimulus intensity.

1.4. *The Common Basis: Weber's Relativity Law*

'In observando discrimine rerum inter se comparatarum non differentiam rerum, sed rationem differentiae ad magnitudinem rerum inter se comparatarum percipimus.'

These are the words by which E.H. WEBER (1834, p. 172)[40] himself generalizes his theory of thresholds. They form the common basis of the approaches of Bernoulli, Fechner, and Stevens, for in all these approaches it is assumed that men face relative rather than absolute changes in stimulus intensities. Equal relative changes are equally perceptible, equally intensive, or equally significant. Whether on the *psychological continuum,* as with Bernoulli and Fechner, equal differences in sensations or, as with Stevens, equal ratios of sensations are perceived as equally significant[41] or whether, as with Weber, there is no functional relationship between stimulus intensity and sensation at all, does not matter very much. *Weber's relativity law* is the common basis of all of the above approaches and it is confirmed by everyday observations. From now on, when the term Weber's law is used in this book, it will refer to this meaning[42].

---

[40] Similarly WEBER (1834, pp. 161 and 173).

[41] Concerning the general interpretation of their laws cf. FECHNER (1860 I, pp. 54–69) and STEVENS (1975, p. 18).

[42] The idea of a more fundamental relativity law underlying the laws of Weber and Fechner was developed by WUNDT (1863, esp. pp. 65–76) and was taken up by WUNDT (1908, esp. pp. 629–645), GROTENFELD (1888), MEINONG (1896), and LIPPS (1902; 1905, pp. 231–287). From a desire to show that Weber's law is compatible with more than just

We can detect an object in both bright and dim light, since the *ratios* of light intensities on the retina are constant, and independently of its distance, because it is the *proportions* of the retina picture that matter and not its absolute magnitude. We perceive a melody independently of the octave in which it is played and independently of the musician's distance from us, since equal frequency *ratios* and equal *ratios* of sound pressure are perceived as equal. Our sensory system has no difficulty in steering a car through daily traffic although, in the course of its evolution, it only had to learn how to make our comparatively poorly equipped bodies function. We live our luxurious lives as matter-of-factly as our ancestors lived their much simpler ones. How could Niels Bohr possibly have been able to explain atomic structure by a planetary model if he had not thought in terms of magnitude *ratios*?

Weber's relativity law is certainly not limited to the mere physiological fact that, for simple physical stimuli, the impulse frequency in nerve fibres is a logarithmic function of stimulus intensity. This is only *one* of its multiple variates. The example of pitch sensation clarifies this point. According to the theory of von Helmholtz (1869), which, after its experimental verification by Galambos and Davis (1943) and its modification by von Békésy (1956), can be considered as valid[43], sound frequency is not, as one might suspect, transformed into a frequency of nerve impulses. Rather, there are specific receptors for sound frequency where the impulse frequency emitted from these receptors has the sole task of transmitting sound pressure, according to a logarithmic law. The fact that an impulse is transmitted by a fiber at all is associated with a particular sound frequency in the central nervous system. But nevertheless, as we know, equal relative changes in sound frequency are perceived as equally significant.

Evidence for a comprehensive relativity law, however, is primarily provided by the empirical investigations into the perception of the length of lines (Stevens, Eisler) or the number of equally dispersed dots on a white card (Thurstone, Guilford). If equal ratios are perceived here as equally significant the central nervous system must be carrying out an

---

the logarithmic function, rather than because the empirical facts required it, these authors proposed the power function of sensation, occasionally even in its special version $\Theta = 1$ (proportionality). The authors seem to have believed that, particularly when equal ratios of stimulus intensities bring about equal ratios of sensation intensities, a 'purely psychological' (Grotenfeld) explanation of the relativity law is needed. Cf. in this context the axiomatic foundation of a comprehensive relation theory given by Krantz (1972).

[43] In 1961 von Békésy was awarded the Nobel Prize for his model of the ear. He rejected Helmholtz's conjecture that the membranes in the cochlea of the ear vibrate according to the frequencies heard. They are unable to vibrate since they are not under tension. Nevertheless, according to von Békésy the membranes are able to percieve specific frequencies, or precisely: ranges of frequencies, just as von Helmholtz had conjectured.

extremely complicated calculation process. The perception procedure must be even more complex if we are concerned with imagined stimuli rather than observed ones. As the success of the number-matching method proves, Weber's relativity law is valid even then.

Thus it turns out that our sensory apparatus is adapted quite generally to relativity. This certainly is not by chance. The reason seems to be that the information embodied in the stimuli produced by our environment are encoded in a ratio language. Equal loudness ratios, equal light-intensity ratios, or equal magitude ratios generally mean equal pieces of information. It seems very plausible that an organism which developed through an evolutionary optimizaton process taking millions and millions of years, indeed millions of generations, has learned to decode the ratio language by using its calculation capacity economically, namely by neglecting the information about the absolute intensities and concentrating instead on their ratios. We should accept this special feature of our perception apparatus as a matter of fact and ask only what it implies for the shape of the von Neumann-Morgenstern function.

2. *Risk Preference and Weber's Relativity Law*

In order to take account of Weber's law the von Neumann-Morgenstern axioms are now extended by the following axiom.

Weak Relativity Axiom: *Equal relative changes in wealth are equally significant to the decision maker.*

The axiom takes up the idea underlying Bernoulli's relativity law but formulates this idea in a way that is suggested by psychophysics. It gives an appropriate description of reality if wealth can be considered as one of the continua in Stevens's system of power functions. This, for example, is the case if, in number-matching experiments for wealth, it can be demonstrated that the numbers people find on their balance sheets and the numbers by which they estimate the magnitude of their wealth are equal or proportional to one another, which is a weak requirement.

It would be wrong to interpret the axiom as postulating that a utility-of-wealth function can be calculated by adding up equal relative changes in wealth. *A fortiori,* it does not require a logarithmic von Neumann-Morgenstern function. On the other hand, of course, the axiom does not exclude a logarithmic, and thus cardinal, utility-of-wealth function. The overwhelming empirical evidence in favor of a logarithmic system of sensation functions underlying Stevens's empirical findings actually suggests such a function. In the multiperiod approach developed in section IV B we shall therefore make use of Fechner's law. In this chapter, however, reference to a logarithmic utility function will only be made for the sake of comparison. The analysis as such does not rely on more than the weak version of the Relativity Axiom presented above.

2.1. *The Relativity Law and the von Neumann-Morgenstern Function*

The question is now which implications can be drawn from the Weak Relativity Axiom for the shape of the von Neumann-Morgenstern function. The measure for the intensity of insurance demand[44], $g = p_{\max} q/E(C)$, defined above can be usefully employed to find an answer. Obviously, the Weak Relativity Axiom implies that the decision problem of an insurance purchaser stays unchanged if his initial wealth a, the possible losses C, and the interest-augmented premium pq all alter by the same percentage, i.e., if $a' = \lambda a$, $C' = \lambda C$, and $p'q = \lambda pq$ for all $\lambda > 0$. This in turn implies that the interest-augmented maximum premium he is willing to pay changes by the same percentage: $p'_{\max} q = \lambda p_{\max} q$. Hence, with

$$(30) \qquad g\lambda^0 = \frac{p'_{\max} q}{E(C')} = \frac{\lambda p_{\max} q}{E(\lambda C)} = \text{const.}, \qquad \lambda > 0,$$

it turns out that the intensity of insurance demand stays constant. In other words, the intensity of demand for an insurance of wealth is independent of the size of wealth. The fact that $p'_{\max} q$ and $E(C')$ are proportional to λ implies that, because $\pi' = p'_{\max} q - E(C')$, the subjective price of risk π' is also proportional to λ. Thus the general version of *Weber's law* brings about what, in the Pratt-Arrow terminology, is called *constant relative risk aversion* or what POLLAK (1970, p. 121) denoted by the term 'weak homogeneity.'

From equation (5) it is already known that the constancy of the intensity of demand for wealth insurance is an implication of the logarithmic utility function favored by Bernoulli. Here the argument is the other way round. Obviously the Relativity Axiom offers the logarithmic utility function as one of the possibilities. The question, however, is whether there are other suitable functions that are also compatible with the Relativity Axiom. Bernoulli's mistake was that he bypassed this question by identifying risk preferences and utility of non-

[44] Cf. equation (II C 17).

random wealth. Equation (II C 14) and the Weak Relativity Axiom imply that $p'_{\max}p = \lambda p_{\max}q = \lambda aq - S(\lambda aq - \lambda C)$. Hence, the complete class of von Neumann-Morgenstern functions that are compatible with a constancy of g is characterized by a linear homogeneity in the certainty equivalent $S(V) = U^{-1}\{E[U(V)]\}$:

$$\text{(31)} \qquad \lambda U^{-1}\{E[U(V)]\} = U^{-1}\{E[U(\lambda V)]\}.$$

This aspect allows us to make use of a theorem by ACZÉL[45] (1966, pp. 151–153) according to which the *only* strictly monotonically increasing[46] functions $U(v)$ that satisfy this requirement are:

$$\text{(32)} \qquad U(v) = \begin{cases} \Theta v^{\Theta}; & \Theta \neq 0, \quad v > 0, \\ \ln v; & v > 0. \end{cases}$$

(Also strictly positive linear transformations are admissible.)

That these functions, which from now on will be called *Weber functions,* are implied by the assumption of constant relative risk aversion has already been shown by PRATT (1964) and ARROW (1965). These authors defined the value of the negative elasticity of marginal utility

$$\text{(33)} \qquad \varepsilon(v) \equiv -\eta_{U'(v),v} = -\frac{U''(v)}{U'(v)}v,$$

that was used above as a measure of curvature[47], as a measure of local relative risk aversion. By using this measure, which in the present case turns out to be constant, the von Neumann-Morgenstern function can be written as

$$\text{(34)} \qquad U(v) = \begin{cases} (1-\varepsilon)v^{(1-\varepsilon)} & \text{for} \quad \varepsilon \neq 1, \\ \ln v & \text{for} \quad \varepsilon = 1. \end{cases}$$

(Note that the previous assumption $g > 1$ implies $\varepsilon > 0$.) For the certainty equivalents we then have

$$\text{(35)} \qquad S(V) = \begin{cases} E(V^{1-\varepsilon})^{1/(1-\varepsilon)}, & \varepsilon \neq 1, \\ \sum_{i=1}^{n} v_i^{w_i}, & \varepsilon = 1, \end{cases}$$

[45] The suggestion for this theorem was given to me by I. Strauß. The theorem has already been used in chapter II D 2.1.2 in connection with the Krelle-Schneider criterion.

[46] This is necessary because of the Axiom of Non-Saturation.

[47] See equation (26).

where the v_i's are the possible end-of-period wealth variates and the w_i's the corresponding probabilities[48]. That these certainty equivalents in fact are linear homogeneous could easily be shown.

It is worth noting that the functions described in (34) include not only the Bernoulli-Fechner function ($\varepsilon = 1$) and Stevens's power function ($\varepsilon < 1$), but also a more curved type ($\varepsilon > 1$). Examples for all these types are illustrated in Figure 5.

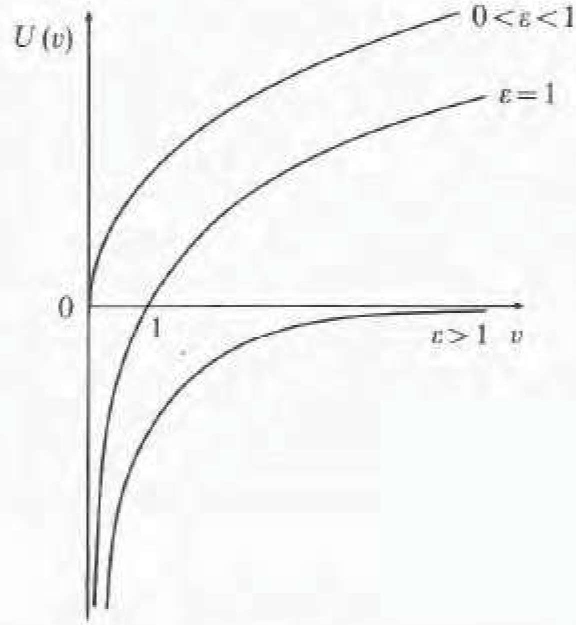


Figure 5

Since we limited our attention to the case of risk aversion, i.e., $\varepsilon > 0$, the figure only shows concave functions. In the case $\varepsilon = 0$ the function $(1-\varepsilon)v^{(1-\varepsilon)}$ describes a ray through the origin and in the case $\varepsilon < 0$ it gives a convex curve that starts horizontally at the origin.

Despite Aczél's theorem, it might nevertheless be thought that there is yet another way of combining the Weak Relativity Axiom with the von Neumann-Morgenstern axioms. What about defining the von Neumann-Morgenstern function over relative (v/aq) rather than over absolute end-of-period wealth[49]? This was the way chosen in principle by Domar and Musgrave (1944, esp. p. 402), Törnqvist (1945, esp. p. 233), and Markowitz (1952b, esp. p. 155). The certainty equivalent

[48] For a continuous density function the certainty equivalent is in the case $\varepsilon = 1$:

$$S(V) = e^{\int_{-\infty}^{+\infty} f(v) \ln v \, dv}.$$

[49] The following argument can equally well be based on the initial wealth not augmented by interest.

of standardized end-of-period wealth, according to approaches of this type, would be $S(V/aq)$ and hence the non-standardized certainty equivalent would be $aqS(V/aq)$. Since the latter expression is a linear homogeneous function of V and aq, the relationship

$$(36) \qquad g\lambda^0 = \frac{p'_{\max} q}{E(C')} = \frac{\lambda p_{\max} q}{E(\lambda C)} = \frac{\lambda aq - \lambda aq S\left(\frac{\lambda V}{\lambda aq}\right)}{\lambda E(C)}$$

would ensure that each arbitrarily given utility function leads to a constancy in the intensity of insurance demand. Even the quadratic function criticized by Hicks[50] would lose the implausible property of increasing absolute risk aversion. Unfortunately, however, these approaches are not admissible since they contradict the Axiom of Ordering. Consider the case of insurance demand and assume that in the beginning of a period the potential insurance buyer receives a gift of amount x/q and is told at the same time that each possible loss, including the 'loss' of size zero, increases by the amount x. If we ask the decision maker for the change in the certainty equivalent of his end-of-period wealth distribution without insurance we get the uncomprehending answer that, obviously, there is no change since the end-of-period wealth distribution is unaffected by these manipulations. The answer satisfies the Axiom of Ordering and requires that

$$(37) \qquad aqS\left(\frac{aq-C}{aq}\right) = (aq+x)S\left(\frac{aq+x-(C+x)}{aq+x}\right).$$

Since in the case $x \neq 0$ this equation obviously can only be satisfied if $S(.)$ is linear homogeneous, we are back to the functions listed in (34) and to these alone! Standardizing the end-of-period wealth distribution thus does not increase the set of von Neumann-Morgenstern functions compatible with the Weak Relativity Axiom[51].

A clear *interpretation* of our preference hypothesis can be obtained by following KRELLE (1968, pp. 144–147)[52] and splitting up $U(v)$ into a utility function $u(v)$ for non-random wealth and a specific risk preference function $\varphi(v)$ such that $U(v) = \varphi[u(v)]$. If the psychophysical

[50] Cf. footnote 17 in chapter II D.

[51] Similar remarks apply to TSIANG's (1972, p. 358) suggestion of adapting the utility function to the decision maker's expected wealth. Cf. our criticism of this suggestion at the end of section II D 2.2.2.

[52] Cf. ch. II C 1.5.

sensation function of wealth is identified with the utility function $u(v)$, then

$$(38) \qquad u(v) = \ln v$$

and hence the specific risk preference function for evaluating probability distributions of *utility* must be

$$(39) \qquad \varphi(u) = \begin{cases} (1-\varepsilon)e^{(1-\varepsilon)u}, & \varepsilon \neq 1, \\ u, & \varepsilon = 1, \end{cases}$$

to ensure that a combination of both functions yields (34)[53].

Since its application by FREUND (1956), the function $(1-\varepsilon)e^{(1-\varepsilon)u}$, $\varepsilon \neq 1$, is known as a von Neumann-Morgenstern function on the objective continuum (wealth) if u is replaced by v. It is convex if $\varepsilon < 1$ and concave if $\varepsilon > 1$. Hence the Weber functions imply either risk aversion or risk loving on the *subjective* continuum depending on whether relative risk aversion on the *objective* continuum exceeds or falls short of unity. In the case $\varepsilon = 1$, the specific risk preference function is linear and so the decision maker is risk neutral on the subjective continuum. The logarithmic function is not modified in this case but its curvature is sufficient to produce risk aversion on the objective continuum.

There is another interesting aspect of Freund's function that can easily be seen by calculating a *certainty equivalent utility* from the approach

$$(40) \qquad (1-\varepsilon)e^{(1-\varepsilon)S[u(V)]} = E[(1-\varepsilon)e^{(1-\varepsilon)u(V)}]$$

such that

$$(41) \qquad S[u(V)] = \frac{\ln E[(1-\varepsilon)e^{(1-\varepsilon)u(V)}] - \ln(1-\varepsilon)}{1-\varepsilon}.$$

---

[53] For example we have

$$(1-\varepsilon)e^{(1-\varepsilon)\ln v} = (1-\varepsilon)(e^{\ln v})^{1-\varepsilon} = (1-\varepsilon)v^{1-\varepsilon}.$$

Note that in the case $\varepsilon \neq 1$ the utility function $u(v)$ is defined up to an additive constant while of course, as we know, $(1-\varepsilon)e^{(1-\varepsilon)u(v)}$ is defined up to a strictly positive linear transformation. It is possible to write

$$\max E[(1-\varepsilon)e^{a+b(1-\varepsilon)u(v)}] = e^a \max E[(1-\varepsilon)e^{b(1-\varepsilon)u(v)}],$$

but in this expression 'b' cannot be taken to the front of the expectation operator. To be able to interpret ε as a measure of absolute risk aversion on the subjective continuum we set $b = 1$.

If here all possible utility levels are increased by the amount x we have

$$S[u(V)+x] = x + S[u(V)], \tag{42}$$

or, equivalently,

$$E[u(V)+x] - S[u(V)+x] = E[u(V)] - S[u(V)]. \tag{43}$$

Equation (43) shows a subjective price of risk expressed in terms of utility before and after the shift in utility. Since this price is obviously independent of the shift, the Weber functions (34) imply not only constant relative risk aversion on the objective continuum but also *constant absolute risk aversion on the subjective continuum*[54].

Despite all formal similarities, Freund's utility function $U(v) = -e^{-\beta v}$ as applied to the objective continuum is not compatible with the Weber functions. Freund's function exhibits constant absolute risk aversion on the objective continuum, that is, a wealth independence of absolute risk aversion[55]. In this respect, it is the opposite of the Weber functions that, as will be spelled out in more detail below, imply a particular wealth dependence of absolute risk aversion and thereby supplement the *subjective* influence on risk evaluation by an *objective* one.

In the next section A 2.2 the implications our preference hypothesis has for the shape of the indifference curves in a (μ, σ) diagram will be examined and, for the sake of comparison, the way Freund's hypothesis appears in this diagram will be considered too. Later, in section A 2.3, there will be an opportunity to investigate the behavioral implications of the two rival hypotheses further.

2.2. *The Relativity Law in the (μ, σ) Diagram*

As we know, the shape of indifference curves in a (μ, σ) diagram cannot be seen independently of an underlying von Neumann-Morgenstern function. Thus the task of this section is to represent the Weber functions listed in (34). Of course the indifference curves exhibit the properties that have already been derived, in particular the slope of zero at the ordinate[56] and, in the case of linear distribution classes, the over-all convexity[57] caused by risk aversion.

---

[54] Cf. equation (II C 5) and the definition of constant absolute risk aversion in chapter II D 2.2.3.

[55] The postulate of wealth independence of risk aversion is the essence of PFANZAGL's (1959a, p. 39; 1959b, p. 288) Consistency Axiom. Hence, Weber's relativity law in connection with Fechner's law implies the validity of the Consistency Axiom on the subjective continuum.

[56] Cf. chapter II D 2.2.3 and II D 2.3.

[57] Cf. chapter II D 2.3.

These properties, as well as others that will be derived in this section, are, however, subject to the constraint that the range over which the utility function is defined includes the ranges of dispersion of the probability distributions considered. This condition implies that the probability distributions are limited to the positive half of the wealth axis, if $\varepsilon < 1$ with and if $\varepsilon \geq 1$ without the origin. For the case of a linear distribution class, it is therefore required that

$$(44) \qquad \mu - \underline{k}\sigma \begin{Bmatrix} > \\ \geq \end{Bmatrix} 0, \quad \text{i.e.,} \quad \mu \begin{Bmatrix} > \\ \geq \end{Bmatrix} \underline{k}\sigma, \quad \text{if} \quad \varepsilon \begin{Bmatrix} \geq \\ < \end{Bmatrix} 1,$$

where $-\underline{k}$ is the highest lower bound of the standardized random variable[58] $Z = (V - \mu)/\sigma$. How the indifference curves are shaped if the probability distributions also extend over the negative half of the wealth axis is discussed in section B.

For *small standard deviations and arbitrary distribution classes* it is easy to calculate the slope of the pseudo indifference curves in the (μ, σ) diagram by referring to equation (II D 51). If, in this equation, the derivative U''' is replaced by

$$(45) \qquad U'''(\mu) = \frac{\varepsilon}{\mu^2} U'(\mu) - \frac{\varepsilon}{\mu} U''(\mu),$$

an expression that follows from a differentiation of the Weber functions (34), then, with a few steps, we reach

$$(46) \qquad \left. \frac{d\mu}{d\sigma} \right|_{U(\mu,\sigma)} \approx \frac{\varepsilon \dfrac{\sigma}{\mu}}{1 + \dfrac{1}{2}\left(\dfrac{\sigma}{\mu}\right)^2 (\varepsilon + \varepsilon^2)}.$$

According to this formula the slope of the pseudo indifference curves is constant as long as the coefficient of variation (σ/μ) of the wealth distribution is constant. Since this is the case on rays through the origin, (46) implies a *homothetic* pseudo indifference-curve system where the single indifference curves can be constructed from one another by a projection through the origin.

Concerning the degree of approximation we may now revert to the examples calculated in section II D 2.2.2. There it was shown that the degree of approximation is a function of the coefficient of variation σ/μ

[58] Cf. Figure 7 in chapter II B and equation (II A 14).

- if the set of linear distribution classes the decision maker thinks possible is independent of the expected level of wealth and
- if the decision maker wants to be able at least to distinguish between distributions whose relative difference in standard deviations exceeds some critical level.

Hence, in the (μ, σ) diagram, points of equal degrees of approximation lie on rays through the origin as is illustrated by the shaded area in Figure 6.

Provided that $\mu > \underline{k}\sigma$, the result of a homothetic indifference-curve system can be confirmed for *large standard deviations under a linear distribution class.* Since (34) gives a marginal utility function of the type

$$(47) \qquad U'(v) = v^{-\varepsilon},$$

which is defined up to a multiplication with a strictly positive constant, equation (II D 60) can be written as

$$(48) \qquad \left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)} = -\frac{E[Z(\mu+\sigma Z)^{-\varepsilon}]}{E[(\mu+\sigma Z)^{-\varepsilon}]} = -\frac{E\left[Z\left(\frac{\mu}{\sigma}+Z\right)^{-\varepsilon}\right]}{E\left[\left(\frac{\mu}{\sigma}+Z\right)^{-\varepsilon}\right]}$$

which again indicates a homothetic system of indifference curves.

The indifference-curve system illustrated in Figure 6 shows the derived properties. Because of the constraint (44), the indifference curves are not plotted below the line $\mu = \underline{k}\sigma$. The way they approach this line is also left open.

A property worth noting is that, for each point above or to the left of the line $\mu = \underline{k}\sigma$, the indifference-curve slope must be smaller than that of the corresponding ray through the origin:

$$(49) \qquad \left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)} < \frac{\mu}{\sigma}.$$

This can easily been shown for continuous density functions $f_z(z; 0, 1)$ if (48) is written in the form

$$(50) \qquad \left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)} = -\int_{-\underline{k}}^{\infty} z\xi(z)\,dz$$

where use is made of the standardized weight factor

$$\xi(z) \equiv \frac{f_z(z;0,1)\left(\frac{\mu}{\sigma}+z\right)^{-\varepsilon}}{\int_{-k}^{\infty} f_z(z;0,1)\left(\frac{\mu}{\sigma}+z\right)^{-\varepsilon} dz}. \tag{51}$$

Clearly (50) implies $d\mu/d\sigma|_U \leq \underline{k}$. Now, for points above the border line $\mu = \underline{k}\sigma$ it holds that $\underline{k} < \mu/\sigma$. Thus (49) is obvious.

Figure 7 is confined to the case of risk aversion ($\varepsilon > 0$). Of course under risk neutrality ($\varepsilon = 0$) and risk loving ($\varepsilon < 0$) the indifference curves would be linear or concave. Rather than studying these irrelevant cases we would do better to find out how the degree of risk aversion affects the shapes of the indifference curves. Differentiating (46) for ε we have

$$\left.\frac{d\left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)}}{d\varepsilon}\right|_{\frac{\sigma}{\mu}} \approx \frac{\frac{\sigma}{\mu}\left[1-\frac{1}{2}\left(\frac{\sigma}{\mu}\varepsilon\right)^2\right]}{\left[1+\frac{1}{2}\left(\frac{\sigma}{\mu}\right)^2(\varepsilon+\varepsilon^2)\right]^2} > 0, \quad \text{if} \quad 0 < \frac{\sigma}{\mu} < \frac{\sqrt{2}}{\varepsilon}. \tag{52}$$

This expression shows that in the case of small coefficients of variation, i.e., when the degree of approximation is particularly good, an increase in the degree of risk aversion raises the indifference-curve slope. This result can be restated by referring to expressions (50) and (51). There, an increase in ε shifts the part

$$\frac{\left(\frac{\mu}{\sigma}+z\right)^{-\varepsilon}}{E\left[\left(\frac{\mu}{\sigma}+Z\right)^{-\varepsilon}\right]}$$

of the weight factor $\xi(z)$ towards lower values of z such that the average z gets smaller and hence $d\mu/d\sigma|_{U(\mu,\sigma)}$ larger. A more precise analysis is given in appendix 1 to this chapter. As expected, the plausible result is

$$\left.\frac{d\left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)}}{d\varepsilon}\right|_{\frac{\sigma}{\mu}} > 0, \quad \frac{\sigma}{\mu} > 0. \tag{53}$$

A homothetic indifference-curve system of the kind described above was postulated by Hicks (1967, p. 114). He called it the standard case 'from which there might be a divergence, in practical experience, in either direction'. Stiglitz (1969a) used it for the sake of comparison and

EBEL (1971, pp. 112 f.) tried to depict the Hicksian postulate by assuming $d\mu/d\sigma|_{U(\mu,\sigma)} = \varepsilon\sigma/\mu$ an expression that approximates our formula (46) for $\sigma/\mu \to 0$[59]. None of the three authors showed its relationship to the von Neumann-Morgenstern function. SCHNEEWEISS (1977a, p. 201 in connection with p. 87) and PYE (1967, p. 115), however, stated that the functions (34) produce a homothetic indifference-curve system. (Pye did not consider $\ln v$.) The homothetic indifference-curve system is implicit in FISHER's (1906, pp. 408 f.) hypothesis that the subjective significance of risk depends on the coefficient of variation (σ/μ) of the wealth distribution. It is implicit also in the certainty equivalent $\mu[1-\alpha(\sigma/\mu)]$ that was used by Palander (1957)[60] and shown by MAGNUSSON (1969, pp. 245–247) to approximate the logarithmic utility function in the case where[61] $\alpha = 1/2$. Apart from these references, however, almost all the rest of the literature making use of the (μ, σ) approach does not refer to the homothetic indifference curve system. In most cases the indifference curves are the concentric circles[62] that can be derived from quadratic utility, although the authors are usually shrewd enough to forgo the plotting of these circles in a diagram.

For a comparison with the indifference curves following from Weber's law (Figure 6), Figure 7 shows those following from the hypothesis of constant absolute risk aversion. This preference hypothesis is implied by FREUND's utility function

$$(54) \qquad U(v) = -e^{-\beta v}, \quad \beta > 0,$$

which, for the subjective continuum, was discussed above in the general form $(1-\varepsilon)e^{(1-\varepsilon)u}$. That Freund's function in turn is the only one compatible with constant absolute risk aversion follows from theorems by PFANZAGL (1959a, pp. 39–41, 55–57; 1959b, pp. 288–292), PRATT (1964, p. 130), and SCHNEEWEISS (1967a, pp. 85–87).

Analogously to (46), for *small standard deviations and arbitrary distribution classes* we have from (II D 51):

$$(55) \qquad \left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)} \approx \frac{\beta\sigma}{1+\dfrac{\beta^2\sigma^2}{2}}.$$

[59] For this reason it is possible to calculate from (46) the approximation

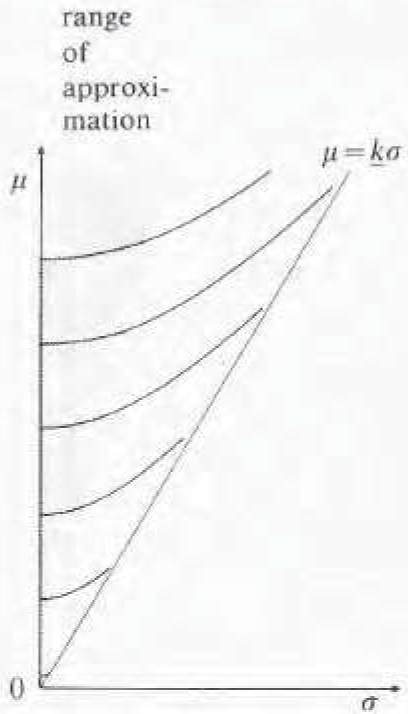
$$S(V) \approx \mu\sqrt{1-\frac{\sigma^2}{\mu^2}}$$

for the certainty equivalent in the case of small dispersions.

[60] Cited according to MAGNUSSON (1969, p. 36).

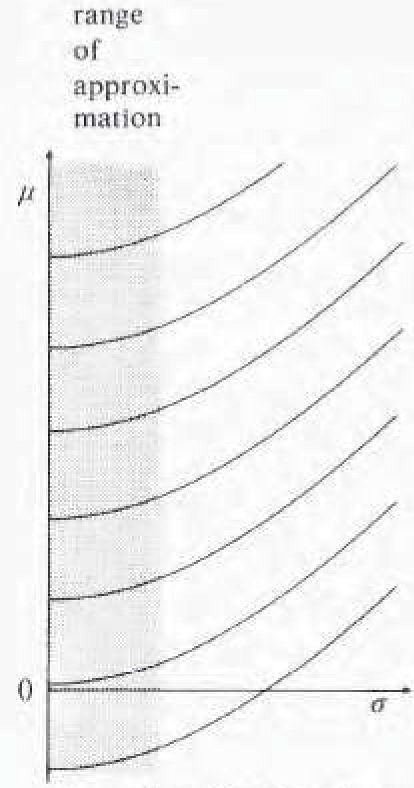
[61] The approximation for the certainty equivalent given in fn. 59 coincides for $\varepsilon = 1$ (logarithmic case) and $\sigma/\mu \to 0$ with the Palander-Magnusson formula.

[62] Cf. chapter II D 2.1.3.



Weber's law

Figure 6



Freund's utility function

Figure 7

Under the same assumptions about the degree of precision and the set of possible linear classes as made above in connection with formula (46) we now infer from the examples calculated in section II D 2.2.2 that points of equal degree of approximation are situated on parallels to the ordinate. In Figure 7 such a parallel is illustrated by the right-hand border of the shaded area.

Analogously to (48), for *large standard deviations under a linear distribution class* we have from (II D 60):

$$(56) \qquad \left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)} = -\frac{E[Z\beta e^{-\beta(\mu+\sigma Z)}]}{E[\beta e^{-\beta(\mu+\sigma Z)}]} = -\frac{E[Z\beta e^{-\beta\sigma Z}]}{E[\beta e^{-\beta\sigma Z}]}.$$

As would be expected, in both cases the level of expected wealth has no influence on the indifference-curve slope. In the case of Freund's function the indifference curves thus can be transformed into one another by parallel shifts along the ordinate.

It can easily be checked that the parameter β in Freund's function (54) coincides with the Pratt-Arrow measure of absolute risk aversion as defined in equation (II D 56). Therefore, in the case of large standard deviations, the result of parallel indifference curves can also be achieved directly from the general formula (II D 64) that relates the wealth dependence of the indifference-curve slopes to the wealth dependence of local absolute risk aversion.

2.3. *Implications for the Intensity of Insurance Demand*

The preference hypothesis derived from Weber's law allows for two factors that influence the subjective price of risk $\pi(V)$ or the intensity of insurance demand[63] $g = [\pi(aq - C) = E(C)]/E(C)$: the decision maker's *subjective risk preference* as measured by the parameter ε and his *objective wealth a*. Thus, in a certain sense, our hypothesis provides a synthesis between Bernoulli's hypothesis according to which *only* wealth explains differences in risk aversion and Freund's hypothesis criticized e.g., by KRELLE (1957, p. 676), where *only* subjective factors are allowed to influence risk aversion.

2.3.1. The Influence of Subjective Risk Aversion

An obvious conjecture can be made concerning this influence[64]. The higher ε, the higher the intensity of insurance demand should be. Indeed this conjecture is correct. According to (52) and (53), when ε rises, the indifference-curve slope gets steeper on any given ray through the origin. Hence the vertical distance π between a point (μ^*, σ^*), $\sigma^* > 0$, and the point where the corresponding indifference curve enters the ordinate rises. Formally, because the indifference-curve slope is a function of the type $s(\mu/\sigma, \varepsilon)$, $s_2 > 0$, we have

$$(57) \qquad \pi = \int_0^{\sigma^*/\mu^*} s(x, \varepsilon)\, dx$$

and hence

$$(58) \qquad \frac{d\pi}{d\varepsilon}, \frac{dg}{d\varepsilon} > 0.$$

It should be noted that this result holds for any given probability distribution that is in the admissible range specified in the beginning of

---

[63] Cf. chapter II C 1.3.

[64] Because of its implications for the subjective price of risk in the case of small dispersions, PRATT (1964) and ARROW (1965) have chosen the parameter ε to measure subjective risk aversion. Cf. the role of the absolute risk aversion measure $\beta \equiv \varepsilon/v$ in equations (II D 55) and (II D 56). That ε has the same relevance for large risks is plausible but not self-evident.

section 2.2 because, for any such probability distribution, a particular indifference-curve system can be constructed. Constraints concerning the class of admissible distributions only become relevant if two genuine distributions are to be compared, they are not relevant if, as in the present case, a genuine distribution is compared with a non-random level of wealth.

2.3.2. The Influence of Wealth

To understand the relationship between wealth and risk evaluation the preference structure following from Weber's law is compared with the one that Freund modelled with (54).

It is known that, according to the Weak Relativity Axiom, a proportional extension *and* shift of the end-of-period wealth distribution leads to a proportional increase in the subjective price of risk. Moreover it is clear that under the hypothesis of constant absolute risk aversion an increase in initial wealth, given the distribution of period income ($\sigma = \text{const.}$), does not affect the price of risk.

The first question is addressed to the *preference structure according to Weber's law.* How does the subjective price of risk (π) change under an increase in initial wealth (a) given the distribution of period income or, more pointedly, how does the intensity of demand for an *insurance of given risk* change if wealth is rising?

The answer, that in a different form has been given by PRATT (1964, pp. 130 f.) and MOSSIN (1968, pp. 555 f.), can easily be found from[65] Figure 8. There the points A', B', and C' are constructed by a projection through the origin from A, B, and C and thus $\pi'/\pi = \overline{OA'}/\overline{OA} = \overline{OB'}/\overline{OB} = \overline{OC'}/\overline{OC}$. Moving, for a given σ, from B to B'' we find that π changes to π''. Of course $\pi'' < \pi'$, but we find in addition that $\pi'' < \pi$. At all points on the curve segment AB, except point A, the slope is higher than at the corresponding points vertically above them in the segment $A'B''$. Hence the integration $\int_0^{\sigma^*} s(\mu/x, \varepsilon)\,dx$ implies $\pi'' < \pi$ or generally

$$\text{(59)} \qquad \frac{d\pi}{da}, \frac{dg}{da} < 0, \quad \sigma > 0.$$

The reason the indifference-curve slope is a falling function of μ is that, according to (47) and (49), points of equal slope are situated on a ray through the origin and that the indifference curves are convex according to the proof given in chapter II D 2.3. Thus Weber's law implies that the

[65] The proof given here is no less general than the one given by Pratt and Mossin, since it holds for any shape of the probability distribution as long as its range is covered by the range over which the utility function is defined. The reason is that B and B' belong to the same linear class.

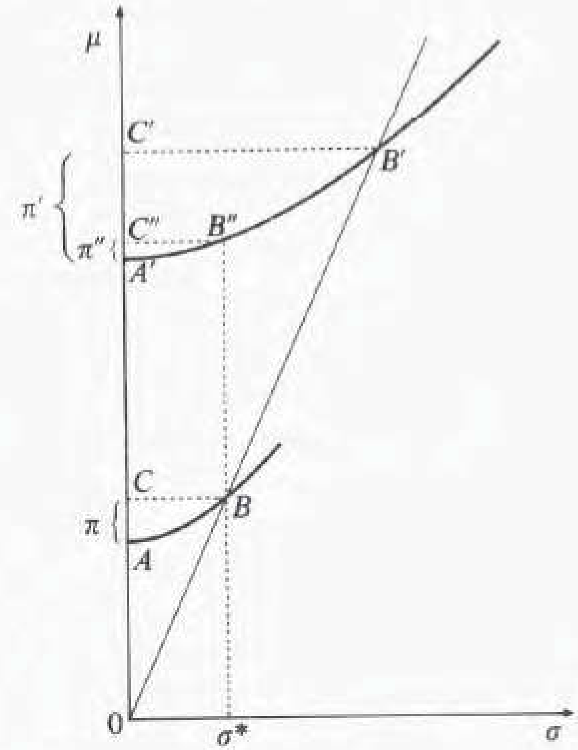


Figure 8

intensity of demand for an insurance of given risk is falling when wealth is rising.

The result shows that the hypothesis of decreasing absolute risk aversion is an implication of Weber's law. In modern literature on risk theory the hypothesis is accepted as being 'supported by everyday observation' (ARROW (1965, p. 35)) and 'intuitively appealing' (BICKSLER (1974, p. 4)). BERNOULLI (1738, § 3 and § 15) had already recognized that decreasing absolute risk aversion is an implication of his logarithmic utility function, and although FISHER (1906, p. 277) did not use the expected-utility approach he, too, argued in favor of this hypothesis. Among insurance practitioners, finally, the hypothesis is generally taken for granted[66]: the fact that nowadays insurance is not bought for umbrellas because *comparatively* small risks are involved is a frequently cited example.

In chapter II D 2.2.3 it was shown that decreasing absolute risk aversion implies a preference for right skewed distributions[67]. It is worth noting that with the present discussion this particular preference can be traced back to Weber's law.

[66] Cf. e.g., FARNY (1961, p. 151). In insurance theory the hypothesis of decreasing absolute risk aversion is not only considered plausible for the insurance purchaser but also for the insurance company. See HELTEN (1973, p. 192).

[67] Cf. in particular the remarks after equation (II D 58).

In their experiments on subjective risk preferences MOSTELLER and NOGEE (1951, pp. 399 f.) observed the, to them disturbing, fact that the experimental subject's risk aversion is dependent on the 'amount of money he has on hand'. This phenomenon, too, is explained by Weber's law[68].

To complete the comparison between the two rival preference hypotheses we now investigate the *hypothesis of constant absolute risk aversion* by asking, how the subjective price of risk changes when the end-of-period wealth distribution undergoes a proportional extension and shift or, alternatively, how the intensity of demand for wealth insurance changes if wealth is rising. The question may easily be answered with the aid of Figure 8, if we interpret an indifference curve as the graph of a function $\pi(\sigma)$ that is defined up to a constant $S(V)$ that measures the level at which this graph enters the ordinate. Obviously, because of the convexity of the indifference curves, the subjective price of risk increases more than proportionally with the standard deviation. Because of the wealth independence of the shapes of the indifference curves this result continues to hold if the standard deviation σ/μ stays constant. Hence, with an increase in wealth, the intensity of demand for *wealth insurance* rises. This conclusion is in striking contrast to the Weak Relativity Axiom and may thus be used for an empirical discrimination between the two rival hypotheses.

2.4. *Result*

Axioms of rational decision making under risk naturally leave substantial scope for differences in individual behavior. This scope follows from the use of general assumptions, but at the same time leads to empty conclusions. Combining Weber's relativity law, which is safely founded in a large body of psychophysical experimental work, with the von Neumann-Morgenstern utility theory, we were able to reduce the scope substantially. A number of interesting conclusions emerge.

Since the von Neumann-Morgenstern function must be such that it implies a linear homogeneity in the certainty equivalent $U^{-1}\{E[U(V)]\}$, only the utility functions

$$U(v) = \begin{cases} (1-\varepsilon)v^{1-\varepsilon}, & \varepsilon \neq 1, \\ \ln v, & \varepsilon = 1, \end{cases}$$

are possible where ε is the absolute value of the elasticity of marginal utility, that is, the Pratt-Arrow measure of the degree of relative risk

[68] MOSTELLER and NOGEE (1951, p. 400) conjectured that the utility function changes with wealth. Cf. the above discussion of equations (36) and (37).

aversion. For almost arbitrary distribution classes but small dispersions, these functions imply a homothetic indifference-curve system in the (μ, σ) diagram. The quality of approximation in this diagram is a function of the coefficient of variation σ/μ of the end-of-period wealth distribution. For distributions from a linear class whose standardized distribution to the left is bounded at $z=-\underline{k}$ there exist indifference curves in the (μ, σ) diagram in the range where $\mu/\sigma > \underline{k}$. These curves are an exact representation of a von Neumann-Morgenstern function: they are homothetic, convex, and enter the ordinate perpendicularly. Important implications for risk evaluation are that the intensity of insurance demand

- rises with risk aversion as measured by ε,
- is independent of wealth in the case of wealth insurance,
- decreases with a rise in wealth if the risk to be insured is given.

Section B
The Bloos Rule

In the preceding analysis it was assumed that the range of dispersion of a probability distribution to be evaluated does not exceed the range over which the Weber functions (A 34) are defined. To avoid the possibility of negative variates of wealth, distributions bounded to the left at $v=\mu-\underline{k}\sigma$ were excluded when $\mu/\sigma<\underline{k}$, $\varepsilon<1$, and when $\mu/\sigma\leq\underline{k}$, $\varepsilon\geq 1$. Moreover, distributions not bounded to the left were generally disregarded.

This exclusion seems very restrictive since among the ones it rules out is the normal distribution which, because of its approximation property for sum variables, has a significant practical relevance. On the other hand, it should not be forgotten that such an approximation, though useful, has its limitations. However similar the distributions that occur in reality seem to be to the normal distribution, in at least one respect there is a significant difference: *actual* wealth cannot become negative, because, quite clearly, no one can lose more than he has. This fact is graphically stated in the phrase 'you can't get blood out of a stone' or, to coin a word, in the 'Bloos rule'. It is true that there are many people who burden themselves with more debt than they can ever hope to repay in their lifetimes, i.e., people whose economic balance sheets, including human capital, indicate negative wealth. However, since the debtor's prison has been abolished, the fact that part of the debt is not redeemable does not worry them[1,2,3].

Let V^n denote the actual or *net* distribution of wealth and let V denote the balance-sheet or *gross* distribution of wealth. Then the BLOOS rule is

$$V^n = \begin{cases} V, & V \geq 0, \\ 0, & V \leq 0. \end{cases} \tag{1}$$

Given the Weber functions this relationship implies a complete preference ordering over gross distributions whose properties will be studied in what follows. It will be useful to carry out this study separately for the cases of weak ($\varepsilon < 1$) and strong ($\varepsilon \leq 1$) risk aversion, since, in the first case the utility function is bounded from below, while in the second it is not.

1. *The Complete Preference Ordering under Weak Risk Aversion ($0 < \varepsilon < 1$): The True Reason for Risk Loving*

1.1. *The Derived Utility Function for Gross Wealth Distributions*

Using (1) and letting the Weber functions (A 34) be denoted by $U^n(.)$ we can construct the following *derived* utility function $U(.)$:

$$U(v) = \begin{cases} U^n(v) = v^{1-\varepsilon}, & v \geq 0, \\ U^n(0) = 0, & v \leq 0, \end{cases} \tag{2}$$

with

$$U'(v) = (1-\varepsilon)v^{-\varepsilon} \quad \text{and} \quad U''(v) = -\varepsilon(1-\varepsilon)v^{-(1+\varepsilon)}, \quad \text{if } v > 0,$$

and with

$$U'(v) = U''(v) = 0, \quad \text{if } v < 0.$$

---

[1] Thus a formal test developed by SCHNEEWEISS (1964; 1967a, pp. 129–160) for finding out the intersection of preference structures that on the class of normal distributions, may just as well be represented in a (μ, σ) diagram as by means of a von Neumann-Morgenstern function cannot be applied. One of the conditions for the application of this test, $\lim_{v \to -\infty} U(v)e^{-\alpha v^2} > -\infty$, $\alpha = \text{const.} > 0$, cf. (1967a, p. 131), is not satisfied.

[2] Recall the definition of wealth given at the beginning of this chapter. According to this definition a complete loss of wealth means that, during the whole of his life, the decision maker still retains enough for subsistence minimum consumption. Because of the limits to attachment usual in countries with a well-developed law system this seems to be a realistic assumption.

[3] The significance of a lower boundary of wealth for the evaluation of risks was also pointed out by SEIDL (1972, pp. 443–445). Seidl did not attempt to integrate this boundary into a fòrmal decision theoretic approach. The following analysis extends the one given in SINN (1982) by considering arbitrary distribution classes rather than binary distributions only.

Since the derived utility function $U(.)$ evaluates the gross distribution V in exactly the same way as the original function $U^n(.)$ evaluates the net distribution V^n we have

$$(3) \qquad E[U(V)] = E[U^n(V^n)]$$

for all elements of the opportunity set.

$U(.)$ is nothing more then an auxiliary mathematical construct that draws all its information from $U^n(.)$. Thus we should not be bothered by the fact that $U(v)$ does not satisfy the Non-Saturation Axiom for $v \leq 0$. $U^n(v^n)$ is compatible with this axiom for all admissible values of v^n and that is sufficient.

By analogy with Figure 10 in section II C 1.2, the following Figure 9 demonstrates the implications of the BLOOS rule for the intensity of insurance demand if, because of $l > aq$, the loss distribution $C = \begin{pmatrix} w & 1-w \\ l & 0 \end{pmatrix}$ is large enough to allow for negative gross wealth, a case that is particularly relevant in the case of liability insurance.

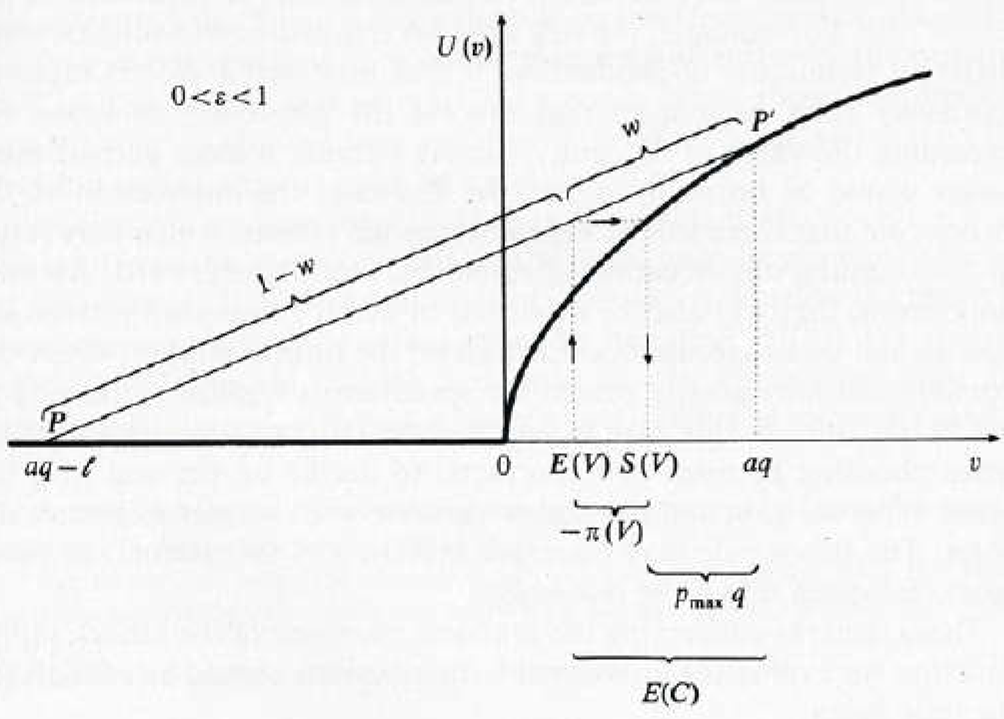


Figure 9

The remarkable aspect of this figure is that the Bloos rule implies a kink in the derived utility function that destroys the over-all concavity that is usually assumed. If the possible loss l is large enough, the certainty equivalent $S(V)$ exceeds the expected level of wealth so that the subjective price of risk $\pi(V)$ is negative and the intensity of insurance demand is smaller than one:

$$(4) \qquad g = \frac{p_{\max} q}{E(C)} < 1.$$

One implication for the case of liability insurance is obvious. Since insurance companies have to demand a premium at least equal to the expected indemnification payment[4], it is preferable for the person facing the liability risk to stay uninsured, although, concerning his subjective preferences, he is a risk averter in the usual sense. The reason is that, even without insurance, the person liable avoids part of the loss which, of necessity, is borne by other parties sustaining the damage. Insurance is not attractive since part of the premium that the buyer pays benefits these other parties and not himself.

Apart from the insurance example that will be considered in more detail later[5], there are a number of other significant implications of the Bloos rule. For example, the rule suggests that, when choosing between different techniques of production, a firm may well decide in favor of extremely risky techniques that involve the possibility of losses far exceeding the value of its equity, simply because a large part of these losses would be borne by others. In this case, the implication of the Bloos rule that there will be negative external effects, which may result in a substantial misallocation of resources, is straightforward. Another implication that will also be discussed in detail[6], shows up in forward speculation when speculators sell short on the futures market, where the possible loss may greatly exceed the speculator's wealth. According to the Bloos rule, in this case it may well be rational for a risk averter, when choosing between two contracts, to decide on the one with the lower expected gain and the higher variance with respect to gross variables. The Bloos rule may therefore explain why speculators, in particular, are often said to be risk lovers.

These remarks concerning the practical relevance of the kinked utility function for evaluating gross wealth distributions should be enough for the time being.

---

[4] Cf. chapter II C 1.2 and 1.3.

[5] In chapter V C 1 and V C 2.3.

[6] In chapter V B 4.

1.2. *Indifference Curves in the* (μ, σ) *Diagram for Linear Distribution Classes*

This section investigates the implications of the BLOOS rule for the shape of the indifference curves in a (μ, σ) diagram in order to facilitate an evaluation of distributions other than the binary distribution considered above. For the sake of simplicity, the analysis is confined to probability distributions that are described by continuous density functions[7]. Moreover we only consider distributions from the same arbitrary linear class[8] with $f_z(z; 0, 1) \equiv f_z(z)$ and $Z \equiv (V-\mu)/\sigma$ or $z \equiv (v-\mu)/\sigma$, respectively, where z is a variate of the standardized random variable Z. For the time being the distributions are assumed to be bounded to the left at $z = -\underline{k}$, $\underline{k} \leq \infty$, and to the right at $z = \bar{k}$, $\bar{k} \leq \infty$. It is assumed that the density is finite except possibly for $z = \bar{k}$.

This characterization of the linear distribution class has the following immediate consequence. Above a ray through the origin $\mu = -\bar{k}\sigma$ (cf. Figure 10) there is a continuum of indifference curves. They all enter the positive part of the μ axis because each distribution, which brings about strictly positive wealth levels with a probability greater than zero, has a strictly positive certainty equivalent. Since the net distribution is bounded at $v = 0$, the lowest conceivable certainty equivalent is zero. Below the border line $\mu = -\bar{k}\sigma$ there is an indifference area. All distributions whose mappings are in this area only extend over the negative half of the wealth axis where $U(v) = \text{const.}$ Thus these distributions bring about a non-random net wealth level of zero and hence the decision maker regards them as equal in value.

Because of the continuity of $U(v)$ in the whole range $-\infty \leq v \leq +\infty$, the indifference-curve slope can, in the usual way, be calculated by use of equation (II D 60)[9]. An interesting question is whether the result of homothetic indifference curves given by equation (III A 48) remains valid. This question can be answered in the affirmative since the equation $U'(\mu + \sigma z) = (1-\varepsilon)\sigma^{-\varepsilon} U'(z + \mu/\sigma)$, that was implicitly used in

---

[7] If there are discrete distributions they may be approximated by a continuous one. Cf. footnote 2 in the introduction to chapter II.

[8] The method of local approximation can no longer be applied since the Weber functions can only be developed into a polynomial down to 50% of the mean. Cf. chapter II D 2.2.2.

[9] The necessary differentiation under the integral

$$\int_{-\infty}^{+\infty} f_z(z) U(\mu + \sigma z)\, dz$$

does not alter the formula despite the discontinuity in the marginal utility function at $v = 0$. Cf. appendix 3 to this chapter where, for another problem, the same mathematical aspect shows up. Formula (5) would cease to hold only if $U(v)$ were discontinuous.

the derivation of (III A 48), not only holds for $\mu+\sigma z>0$, but also for $\mu+\sigma z<0$ since in this case, from (2), $U'(\mu+\sigma z)=0$. Thus the formula

$$(5)\qquad \left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)} = -\frac{E\left[ZU'\left(\frac{\mu}{\sigma}+Z\right)\right]}{E\left[U'\left(\frac{\mu}{\sigma}+Z\right)\right]}$$

remains valid in the case of gross wealth distributions with negative variates so that the indifference-curve system is also homothetic in the range below or to the right of the curve $\mu=\underline{k}\sigma$ (cf. Figures 6 and 10).

Over some range, the indifference curves plotted in Figure 10 have negative slopes, which is a sign of risk loving behavior. This property is already necessitated by the fact that *all* distributions with $\mu/\sigma>-\bar{k}$ have strictly positive certainty equivalents and is plausible in the light of the convexity of the utility function brought about by the BLOOS rule.

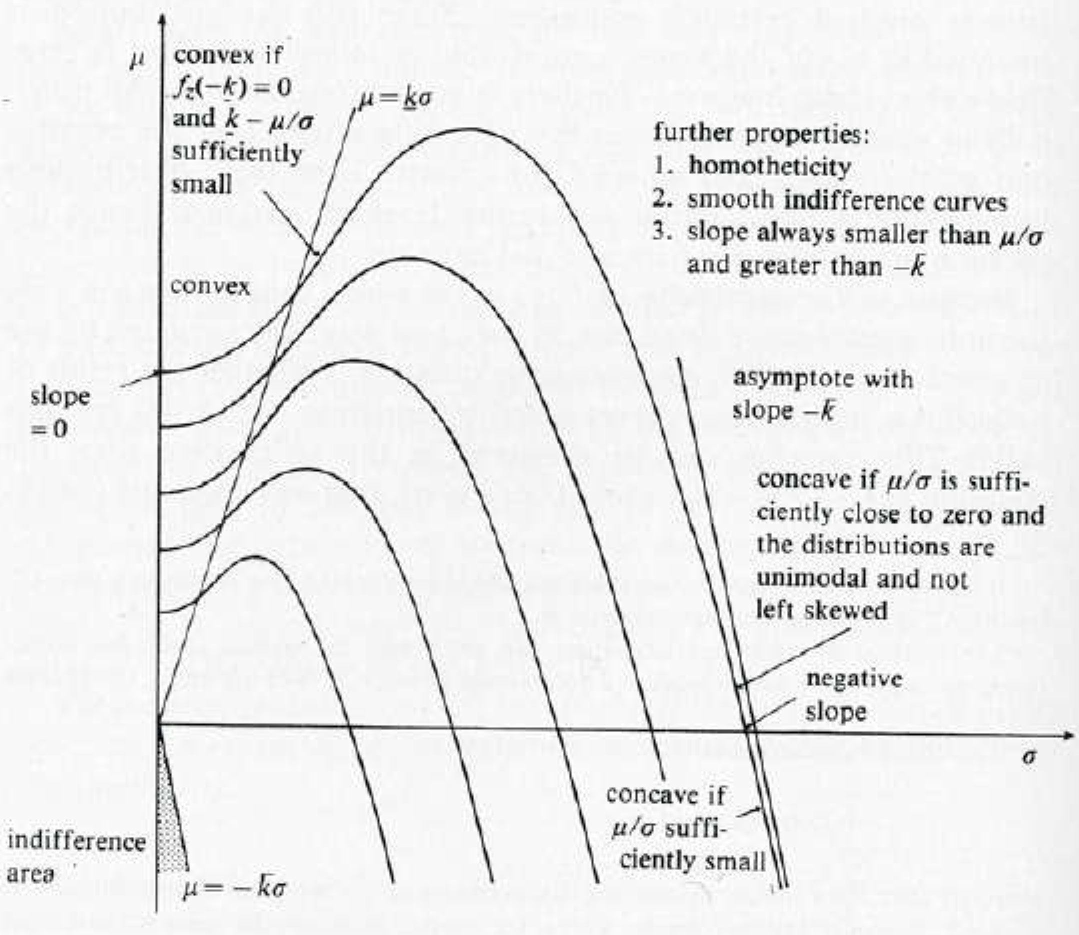


Figure 10

More precise information on the indifference-curve slope when μ/σ is in the range $-\bar{k} < \mu/\sigma \le \underline{k}$ can be gained by inspection of (5). According to this expression, the slope is the negative of a weighted average of the possible variates of Z where the weights are $f_z(z)U'(z+\mu/\sigma)/E[U'(Z+\mu/\sigma)]$. Obviously zero weights are attached to all $z > \bar{k}$, since $f_z = 0$, and to all $z < -\mu/\sigma$, since $U'(v) = 0$ for $v < 0$. Hence $-\mu/\sigma \le E(ZU')/E(U') \le \bar{k}$ and thus $-\bar{k} \le d\mu/d\sigma|_{U(\mu,\sigma)} \le \mu/\sigma$. Although this information about the indifference-curve slope is rather limited, it confirms the fact that, with $-\bar{k} < \mu/\sigma < 0$, there is a range where the slope is negative.

To obtain further information equation (5) is written in the explicit form

$$\text{(6)} \qquad \left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)} = -\frac{\int_{-\mu/\sigma}^{\infty} z f_z(z)(1-\varepsilon)\left(\frac{\mu}{\sigma}+z\right)^{-\varepsilon} dz}{\int_{-\mu/\sigma}^{\infty} f_z(z)(1-\varepsilon)\left(\frac{\mu}{\sigma}+z\right)^{-\varepsilon} dz}.$$

Since[10] $\lim_{z\to-\mu/\sigma+}(z+\mu/\sigma)^{-\varepsilon} = \infty$ it is tempting to conjecture that the weight for $z = -\mu/\sigma$ is dominating all others so that $d\mu/d\sigma|_{U(\mu,\sigma)} = \mu/\sigma$ if $f_z(-\mu/\sigma) > 0$. In the range $-\bar{k} < \mu/\sigma < \underline{k}$, the indifference-curve system would then be a set of rays through the origin. Appendix 2 shows this conjecture to be wrong. (Substitute[11] $A \equiv d\mu/d\sigma|_{U(\mu,\sigma)}$, $\mu/\sigma \equiv y$, $z \equiv w$, $\varepsilon \equiv \Theta$.) As long as the assumption $0 < \varepsilon < 1$ that underlies this section is maintained we have

$$\text{(7)} \qquad \left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)} < \frac{\mu}{\sigma}, \quad \text{if} \quad \frac{\mu}{\sigma} > -\bar{k} \quad \text{and} \quad \infty > f_z\left(-\frac{\mu}{\sigma}\right) > 0.$$

This result will hold *a fortiori* if $f_z(-\mu/\sigma) = 0$, perhaps because the considered type of distribution is multimodal or because[12] $\mu/\sigma > \underline{k}$, the situation analyzed in section A.

It is clear that the lower boundary $-\bar{k}$ of the slope will never be reached when variates of Z in the range $-\mu/\sigma \le z < \bar{k}$ occur with strictly positive probability. If, however, $\mu/\sigma \to -\bar{k}$ then the range of z values used for calculating $E(ZU')$ becomes even narrower. So we must

---

[10] To distinguish limits from above and from below '+' or '−' are placed after the variable denoting the limit.

[11] In parts of the appendices different symbols are used since the mathematical problems treated are considered in various contexts.

[12] This was shown in equations (A 49)–(A 51).

generally conclude that[13]

$$(8)\qquad \left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)} > -\bar{k}, \quad \lim_{\mu/\sigma \to -\bar{k}} \left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)} = -\bar{k}.$$

Hence, for μ/σ sufficiently small, the indifference curves are nearly parallel to the lower boundary $\mu = -\bar{k}\sigma$ of the range where indifference curves exist.

In addition to the information about the slope of the indifference curves, information about the curvature may be of interest. From Tobin's proof presented under point (2) in section II D 2.3 we know that, in the range $\mu/\sigma > \underline{k}$, the indifference curves are strictly convex. To find out about the curvature in the range $-\bar{k} \leq \mu/\sigma \leq \underline{k}$ we differentiate (5) with respect to μ/σ. The result of this calculation, carried out in appendix 3, is:

$$(9)\qquad \frac{d \left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)}}{d\frac{\mu}{\sigma}} = \frac{\varepsilon(1-\varepsilon)}{\beta} \int_{-\mu/\sigma}^{\infty} \left[f_z\left(-\frac{\mu}{\sigma}\right) - (1-\Gamma) f_z(z) \right] \left(\frac{\mu}{\sigma} + z\right)^{-(1+\varepsilon)} dz \left(\frac{\mu}{\sigma} - \left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)}\right),$$

$$1 > \Gamma > 0, \quad \text{if} \quad f_z\left(-\frac{\mu}{\sigma}\right) = 0,$$

$$\Gamma = 0, \quad \text{if} \quad f_z\left(-\frac{\mu}{\sigma}\right) > 0,$$

$$\infty > \beta > 0.$$

Since β, $\varepsilon(1-\varepsilon)$, and, because of (7), the terms in brackets behind the integral are strictly positive and finite, so we only have to consider the integral itself.

As shown in appendix 4, it is finite if, as assumed, $0 < \varepsilon < 1$. (Substitute the integral for A and set $\mu/\sigma \equiv y$, $z \equiv w$, $f_z(-\mu/\sigma) - (1-\Gamma) f_z(z) \equiv$

[13] In the case of discrete probability distributions rather than continuous ones that can be described by density functions, the slope may take on the value $-\bar{k}$ even for $\mu/\sigma > -\bar{k}$. The sufficient condition for this case is that the variate $z = -\bar{k}$ obtains with positive probability and that further variates in the range $-\mu/\sigma \leq z < \bar{k}$ are impossible. For a binary distribution, this means that in the *whole* range below the line $\mu = \underline{k}\sigma$ the indifference curves have a slope of size $-\underline{k}$. The discrete distributions and the corresponding indifference-curve systems can be approximated by the use of continuous distributions as closely as we wish.

$f_w(w)$, $(1+\varepsilon) \equiv \Theta$.) Since the utility curve is kinked at $v=0$ the suspicion could arise that the indifference curves are also kinked somewhere, at least on the border line $\mu = \underline{k}\sigma$. The finiteness of the integral allays this suspicion. The indifference curves are smooth everywhere.

Whether the indifference curves are convex, concave, or linear is determined by the sign of the integral in (9). In the case $\mu/\sigma > \underline{k}$ it is negative because of $f_z(-\mu/\sigma) = 0$ and $\Gamma < 1$; this restates the convexity proved by Tobin. It is worth noting that the sign of the integral stays negative even in the case $\underline{k} - \mu/\sigma > 0$, provided that this difference is sufficiently small and provided that the density function has the property $f_z(-\underline{k}+) = f_z(-\underline{k}-)$. Thus the indifference curves stay convex in the neighborhood of the line $\mu = \underline{k}\sigma$ even for $\mu/\sigma < \underline{k}$. If, however, the density function is 'truncated' so that $f_z(-\underline{k}+) > 0$ while $f_z(-\underline{k}-) = 0$, then immediately below this line there will be a concave segment provided $f_z(-\underline{k}+)$ is sufficiently large.

For a unimodal distribution class the integral is strictly positive if the mode (M) is zero or negative for, in this case, $f_z(-M/\sigma) - f_z(z) > 0$ $\forall z > -M/\sigma$. Hence, for such a class, the indifference curves are definitely concave if M/σ is negative or sufficiently close to zero. In the case of a right skewed or a symmetrical distribution where $M \leq \mu$, this also means that there exists some $x = \text{const.} > 0$ such that concavity is ensured for all $\mu/\sigma < x$. In the case of multimodal distributions, the indifference curves may consist of various convex and concave segments. In this case, concavity is only ensured when μ/σ is small enough so that the highest mode is close enough to zero.

The simplest version of an indifference curve system that can normally be expected is shown in Figure 10. The properties that have been derived are labelled.

Up to now, only distributions that are bounded to the left have been considered. This does not seem unrealistic. In many practical problems even the gross distribution of wealth in the sense of a balance sheet item appears to have this property since the maximum loss is often limited to capital ventured in a particular enterprise rather than to the decision maker's personal wealth. We should think here, for example, of forms of speculation that tie up capital, of share holding, or of the participation in other limited-liability enterprises. On the other hand, there are a number of decision problems like speculation by selling short or liability insurance, where wealth distributions that disperse very widely to the left have to be evaluated. These problems legitimate the consideration of the limiting case of distributions that are unbounded to the left. In addition, of course, the normal distribution creates some interest in this case, although it must be admitted that there are hardly any problems where the normal distribution approximates the left tails of the appropriate distributions particularly well.

For $\underline{k}=\infty$, the area of indifference curves where all variates of V are positive no longer exists. For any $\sigma>0$ the BLOOS rule affects the indifference-curve slope. Clearly this aspect does not change those particular conclusions, derived above for $\mu/\sigma<\underline{k}$, that were not confined to the neighborhood of the $\mu=\underline{k}\sigma$ line. Thus, the property of homotheticity will continue to hold, and the indifference curves will be negatively sloped and concave for μ/σ sufficiently small. The question remains, however, of how the indifference curves are shaped in the neighborhood of the ordinate. Are they still approaching the ordinate perpendicularly, and if so, will there still be some range of convex indifference curves in the neighborhood of the ordinate where the decision maker behaves as a risk averter?

The first part of this question can easily be answered. Since, in the present case $0<\varepsilon<1$, numerator and denominator of (5) are finite[14], the discontinuous region of $U'(v)$ at $v=0$ can be approximated by a continuous marginal utility function as closely as we wish. In the limit as $\sigma\to 0$, for a continuous marginal utility function, the numerator takes on the value $E(Z)U'(\mu)=0$ and the denominator the value $U'(\mu)$[15]; moreover the weight of some given range of approximation around $v=0$ vanishes[16]. Thus, as before, the indifference curves are horizontal at the ordinate:

$$(10)\qquad \lim_{\sigma\to 0}\left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)}=0.$$

[14] This followed, e.g., from the calculation of equation (6) in appendix 2.

[15] Cf. SCHNEEWEISS (1967a, pp. 128 f.).

[16] That removing the discontinuity has a negligible influence can be shown by using expression A in appendix 2. First, within equation (5) from the text above, the substitutions $y\equiv\mu/\sigma$, $z\equiv w$, $\varepsilon\equiv\Theta$ are carried out. Then, the shape of the marginal utility curve is made continuous in the interval from $-y$ through $-y+\Delta$ by the use of a function $\chi(w)$ where

$$\int_{-y}^{-y+\Delta} f_w(w)\chi(x+w)\,dw=\int_{-y}^{-y+\Delta} f_w(w)(x+w)^{-\Theta}dw$$

is assumed, so that γ in equation (9) of appendix 2 remains unchanged. This modification changes the value of $\lim_{x\to y+}\alpha$ by a finite amount. If before the modification Δ was chosen so as to ensure that γ is sufficiently close to unity, then, independently of this modification, equations (11) and (3) of the appendix imply $A\approx\lim_{x\to y+}\beta$ where each desired degree of approximation can be reached. But even if a very high degree of approximation is not desired, the subsequent taking of the limit $y\to\infty(\sigma\to 0)$ implies that $\lim_{x\to y+}\gamma$ approaches unity and hence the question of whether or not the marginal utility curve is continuous in the range from $-y$ through $-y+\Delta$ turns out to be irrelevant as long as $0<\Theta<1$. If $\Theta\geq 1$ then, because of equation (21) from the appendix, we always have $\lim_{x\to y+}\gamma=0$ so that the above reasoning is no longer valid.

The second question can be answered by reference to expression (9). Obviously convexity prevails, despite the Bloos rule in operation, if, in the limit as $\mu/\sigma \to \infty$, the integral ($\int$) in (9) approaches a strictly negative value. To find out whether this is the case, some formal analysis seems necessary.

Assume that the prevailing linear distribution class has the property

$$(11) \qquad f_z(z) > 0, \quad f_z'(z) \geq 0, \quad \text{for} \quad z \leq -\varrho,$$

where ϱ is some constant that is chosen sufficiently large, and let $\mu/\sigma = \varrho + x + y$, $0 < x < \infty$, $0 < y < \infty$. Moreover, write the integral in (9) in the form

$$(12) \qquad \int = \int_{-\varrho-x-y}^{-\varrho} [f_z(-\varrho-x-y) - f_z(z)](\varrho+x+y+z)^{-(1+\varepsilon)}\,dz$$
$$+ f_z(-\varrho-x-y)\int_{-\varrho}^{\infty} (\varrho+x+y+z)^{-(1+\varepsilon)}\,dz$$
$$- \int_{-\varrho}^{\infty} f_z(z)(\varrho+x+y+z)^{-(1+\varepsilon)}\,dz.$$

Then it is possible to derive a sufficient condition for the sign being negative in the limit as $\mu/\sigma \to \infty$. Obviously, by construction, it holds that $\int_{-\varrho-x-y}^{-\varrho} \ldots dz \leq 0$. Hence $\int$ is smaller than or equal to the sum of the other two items on the right-hand side of (12).

Consider now the inequality

$$(13) \qquad \left(\frac{x+y}{x}\right)^{-(1+\varepsilon)} \int_{-\varrho}^{\infty} f_z(z)(\varrho+x+z)^{-(1+\varepsilon)}\,dz$$
$$< \int_{-\varrho}^{\infty} f_z(z)(\varrho+x+y+z)^{-(1+\varepsilon)}\,dz$$

which follows from the facts that

$$(14) \qquad \frac{x+y}{x}(\varrho+x+z) = \varrho+x+y+z, \quad \text{if} \quad z = -\varrho,$$

$$(15) \qquad \frac{x+y}{x}(\varrho+x+z) > \varrho+x+y+z, \quad \text{if} \quad z > -\varrho,$$

and that $(.)^{-(1+\varepsilon)}$ is a strictly decreasing function. Utilizing (13), we clearly have

$$(16) \qquad \int < f_z(-\varrho-x-y)\int_{-\varrho}^{\infty} (\varrho+x+y+z)^{-(1+\varepsilon)}\,dz$$
$$- \left(\frac{x+y}{x}\right)^{-(1+\varepsilon)} \int_{-\varrho}^{\infty} f_z(z)(\varrho+x+z)^{-(1+\varepsilon)}\,dz.$$

Thus $\lim_{\mu/\sigma\to\infty} \int < 0$ if the right-hand side of (16) becomes negative as $y\to\infty$. This in turn is the case, if the first item on the right-hand side vanishes 'faster' than the absolute value of the second. Since $\frac{\partial}{\partial y}\int_{-\varrho}^{\infty}(\varrho+x+y+z)^{-(1+\varepsilon)}\,dz<0$ the second term on the right-hand side of (16) will definitely dominate the first if

$$\text{(17)} \qquad \lim_{y\to\infty}\frac{f_z(-\varrho-x-y)}{(x+y)^{-(1+\varepsilon)}}=0.$$

Equation (17) therefore is a sufficient condition for the indifference curves being convex in the neighborhood of the ordinate.

As an example, consider the normal distribution where $f_z(z)=(1/\sqrt{2\pi})e^{-z^2/2}$. Here, with $z\to-\infty$, the density vanishes faster than $e^{-|z|/2}$ or, equivalenty, with $z=-(\varrho+x+y)$ and $y\to\infty$ faster than $e^{-|-\varrho-x-y|/2}=e^{-(\varrho+x)/2}e^{-y/2}$. Since

$$e^{-y/2}/(x+y)^{-(1+\varepsilon)}=e^{-y/2}/e^{-(1+\varepsilon)\ln(x+y)}=e^{(1+\varepsilon)\ln(x+y)-y/2}$$

this implies that the normal distribution meets condition (17), provided that $\lim_{y\to\infty}[(1+\varepsilon)\ln(x+y)-y/2]=-\infty$. Obviously this is the case. Hence, at least for the normal distribution and all distributions whose densities on the left-hand side converge faster, the indifference curves are convex if, given μ, σ is chosen sufficiently small.

As a final problem in the analysis of distributions that extend over the negative half of the wealth axis the role of the risk aversion parameter ε should be briefly considered. Here the result

$$\text{(18)} \qquad \left.\frac{d\left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)}}{d\varepsilon}\right|_{\frac{\sigma}{\mu}}>0$$

that was previously achieved with (A 48) still prevails, since the reasoning of appendix 1 is completely unchanged.

Thus we may summarize as follows. Although the subjective preferences of the decision maker exhibit risk aversion, albeit moderate because of $\varepsilon<1$, the BLOOS rule produces risk loving behavior provided the gross distribution extends widely enough over the negative half of wealth axis.

In the case of a linear distribution class bounded to the left at $\mu-\underline{k}\sigma$, $\underline{k}<\infty$, the indifference-curve system in the (μ,σ) diagram has the usual properties in the range $\mu/\sigma>\underline{k}$, but, in the range $\mu/\sigma<\underline{k}$, the indifference curves are negatively sloped and concave for μ/σ sufficiently small. With $\bar{k}$ as the upper boundary of the standardized random variable characterizing the distribution class, in the limit as μ/σ approaches $-\bar{k}$, the indifference-curve slope also approaches this value.

In the case of a linear distribution class unbounded to the left, despite the Bloos rule in operation, the indifference curves have the usual properties in the neighborhood of the ordinate provided that, for $z \to -\infty$, the density converges at least as fast as that of a normal distribution. For sufficiently small values of μ/σ the question of a boundedness to the left is irrelevant for the shape of the indifference curves. As with all distributions unbounded to the right, with the normal distribution the indifference curves approach vertical asymptotes for μ/σ sufficiently small.

At any point in the (μ, σ) diagram the indifference-curve slope is a rising function of the measure of relative risk aversion ε.

The indifference-curve system is homothetic.

1.3. *Critique of the Subjectivist Foundation of Risk Preference*

The explanation of risk loving behavior as given by the Bloos rule is at variance with traditional explanations of this behavior. It does not have very much in common with a subjective inclination towards risk except for the fact the utility function has to be bounded for $v \to 0$. Figure 11 compares the utility curve (2) to the classical curves[17] suggested by Törnqvist (1945), Friedman and Savage (1948), and Markowitz (1952b).

A common feature of the classical proposals is that the convex segments in the medium ranges of the curves are explained by the empirical observation that, despite negative expected net gains, people participate in gambling. Markowitz and Törnqvist place the convex segment to the right of the initial wealth a since the range of gambling seems to be there. This assumption implies that there is no given utility-of-wealth curve but that the positions of the curves are dependent on initial wealth. Friedman and Savage argue that the convex segment should be at medium levels of wealth since comparatively poor people seem to be particularly interested in gambling[18].

[17] Friedman and Savage call the argument of their utility curve 'income'. Their reasoning, however, describes a utility-of-wealth function better. Törnqvist and Markowitz assume that the position of the utility curve depends on the decision maker's initial wealth (a).

[18] The authors unanimously explain the concave curve segments to the left of the convex ranges with the preference for buying insurance. The right-hand concave segment is explained by Friedman and Savage with the argument that people seem to have a preference for diversifying prizes in gambling. From a more formal point of view Markowitz and Törnqvist, however, argue that the utility function is bounded from above to avoid the St. Petersburg Paradox. With a similar argument Markowitz finally legitimates a lower boundary to utility which produces a convex segment.

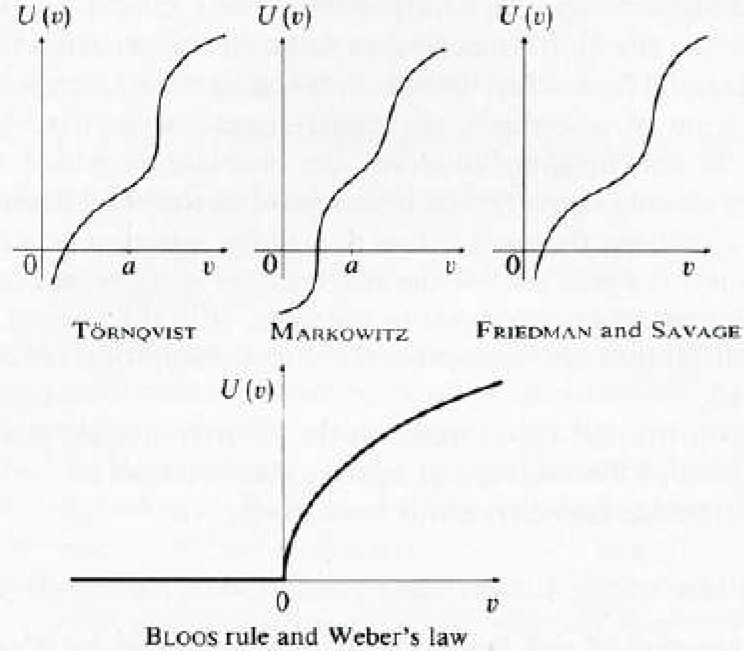


Figure 11

The classical arguments for a convex segment in the utility function are empirically doubtful, inconsistent, and imply irrational behavior under risk. If Friedman and Savage are right, then the intensity of insurance demand must be very low for people from the upper middle class, and no one from this class would want to play roulette[19]. These implausible implications could be removed by adopting the Törnqvist-Markowitz hypothesis that the total curve shifts with an increase in wealth. However, as we know from the discussion of the Weak Relativity Axiom, this solution demands the high price of violating the Axiom of Ordering[20].

These problems suggest that the idea as such of deriving the shape of the utility curve from gambling behavior is misleading. The shortcomings of this type of reasoning are very obvious in the light of some peculiarities of gambling that are hardly compatible with the von Neumann-Morgenstern axioms.

- First, it is necessary to mention the fun of observing complicated game procedures which, as we know, is incompatible with the Axiom

[19] This comment also applies to HAKANSSON (1970b), MASSON (1972), and APPELBAUM and KATZ (1981) who derive the Friedman-Savage utility function from an intertemporal optimization approach with capital-market imperfections. For an extensive criticism of the Friedman-Savage utility function see BAILEY, OLSON, and WONNACOTT (1980).

[20] Cf. in section A 2.1 the discussion centering around equations (A 36) and (A 37).

of Ordering[21]. This fun may explain why people participate in games of chance, but it does not imply anything for the shape of the utility function for serious economic decision making.

- Moreover there is a good case to be made for ALLAIS's (1952, p. 132) explanation of gambling, namely, that the stake is below a subjective threshold while the prizes are beyond it. Again this argument has no bearing on the kinds of economic decision making we are considering in this book. Thresholds do not seem to be important in insurance demand, portfolio choice, or speculation.
- A related argument is that people are inclined to overestimate small, but underestimate large, probabilities. This argument may also explain gambling since the probabilities of winning are usually very small. At any rate, YAARI (1965) felt that an explanation of this sort was needed since he was unable to detect the convex segment in the utility function in his experimental work on risk preferences.
- Finally, doubts must in principle be raised about applying the von Neumann-Morgenstern theory of rational behavior under risk. The attempt by gamblers to outwit probability theory by their 'crystal ball' strategies is surely a good example of irrational behavior. It seems there is a good deal of wisdom in what HICKS (1962, p. 793) says[22] when comparing gambling with portfolio decisions: 'They [the portfolio decisions] are work; gambling is relaxation. To expect consistency in gambling is futile, for gambling is a rest from consistency.'

Rather than trying to explain gambling behavior, it would be better to try explaining types of risk loving behavior that have nothing to do with the fun of gambling or thresholds in perception. Such behavior is to be observed among people who are clever enough to discover what their optimal decisions are.

Why is it that most people obviously have such a low preference for automobile liability insurance that governments had to make this type of insurance compulsory? Why is the entrepreneur, who is up to his ears in debt, willing to risk everything on one more attempt? Why do shipowners build their tankers like tin cans that break open at the slightest impact and spew their oil into the sea, causing losses far greater than the value of the tanker and its contents put together? In all such cases, there

---

[21] If the fun of gambling is independent of stakes and prizes and is merely added to expected utility then MARKOWITZ (1952b, pp. 157 f.) may be right in saying that a concave utility function implies that '... when millionaires play together, they play for pennies'. If however, the completely implausible assumption of independent utilities is removed then it may well be possible that millionaires play with large stakes, even though in serious and less pleasant economic decision problems their behavior exhibits risk aversion.

[22] Cf. also HICKS (1931, p. 181).

is a particular lack of concern for very large negative variates of the wealth distribution. The BLOOS rule, as reflected in the derived utility function, offers an explanation.

2. *The Complete Indifference-Curve System in the Case of Strong Risk Aversion* ($\varepsilon \geq 1$)*: The Implicit Lexicographic Ordering*

In contrast to the case $\varepsilon < 1$, it is now impossible to derive a complete von Neumann-Morgenstern utility function for gross wealth similar to (2), for $\varepsilon \geq 1$ means that the Weber functions (A 34) are unbounded at[23] $\upsilon = 0$: $\lim_{\upsilon \to 0} U^n(\upsilon) = -\infty$. Nevertheless the BLOOS rule remains valid. It is a matter of indifference whether a person loses only his wealth or whether in addition, he is burdened with a debt that he can never repay. Either situation is a complete disaster.

We must thus conclude that at $\upsilon = \bar{\upsilon} = 0$ there is a *lexicographic* critical wealth level so that a maximization of the survival probability becomes the predominant aim:

$$\text{(19)} \qquad \max W(\upsilon > 0).$$

The fact that a combination of Weber's law and the BLOOS rule renders *possible* a lexicographic level of wealth just where $\bar{\upsilon} = 0$, is compatible with the general discussion of the theory of lexicograhic preferences given in chapter II B. In section 1.2 of that chapter we found that a lexicographic critical wealth level, if it exists, is situated at $\bar{\upsilon} = 0$.

Given the information (19), for a linear distribution class it is possible to construct pseudo indifference curves in the (μ, σ) diagram. Since the geometrical locus of points with equal survival probability is defined by the condition[24]

$$\text{(20)} \qquad \frac{\mu - \bar{\upsilon}}{\sigma} = \text{const.},$$

the pseudo indifference curves are rays through the origin; this is illustrated in Figure 12.

However, the total area in the (μ, σ) diagram is not filled with pseudo indifference curves, for the lower ($\underline{k}$) and upper ($\bar{k}$) boundaries of the

---

[23] This property implies a constraint on the range where the Archimedes Axiom is valid. The problem is taken up in the following section C 2.

[24] Cf. equations (II B 5) and (II B 6).

standardized distribution $Z = (V - \mu)/\sigma$ appear on the scene. This, too, is shown by Figure 12.

From below, the area of pseudo indifference curves is bounded by the ray through the origin $\mu = -\bar{k}\sigma$ below which there is an indifference area[25]. The curve is of the same kind as that depicted in Figure 10 and hence we do not have to elaborate upon it.

More important is the upper boundary $\mu = \underline{k}\sigma$, above which, in the case of wealth distributions bounded to the left, there is the range of substitutive indifference curves well known from Figure 6. If a choice has to be made between distributions from this range, then of course the predominant aim of maximizing the survival probability is irrelevant since all of these distributions ensure survival.

In the discussion of Figure 6 the question of how the indifference curves are shaped in the neighborhood of the curve $\mu = \underline{k}\sigma$ was left open. This question will now be considered so that the areas of normal and pseudo indifference curves can be combined without a break. For the case of a bounded utility function ($0 < \varepsilon < 1$), it was shown that (cf. equation (7)) $d\mu/d\sigma|_{U(\mu,\sigma)} < \mu/\sigma$ so that the indifference curves approach the line $\mu = \underline{k}\sigma$ with a slope lower than $\underline{k}$. In the present case $\varepsilon \geq 1$ such a possibility is not excluded. Appendix 2, particularly with equations (25)–(29), shows that, for density functions that at the left continuously approach 0, i.e., that are characterized by $f_z(-\underline{k}+) = f_z(-\underline{k}-) = 0$, there are the following possibilities:

$$(21) \qquad \lim_{\mu/\sigma \to \underline{k}+} \left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)} \begin{cases} = \mu/\sigma & \text{if} \quad \varepsilon \geq 2, \\ < \mu/\sigma & \text{if} \quad \varepsilon < 2. \end{cases}$$

(When using the appendix consider only the calculation of $\lim_{x \to y+} B$, substitute $d\mu/d\sigma|_{U(\mu,\sigma)}$ according to (A 48) for B and set $y \equiv \underline{k}$, $x \equiv \mu/\sigma$, $\Theta \equiv \varepsilon$, and $w \equiv z$.)

Thus, in the case of a comparatively weak risk aversion ($\varepsilon < 2$), it stays true that the indifference curves approach the curve $\mu = \underline{k}\sigma$ at an angle. But if a stronger risk aversion ($\varepsilon \geq 2$) prevails, as is assumed in Figure 12, then the indifference curves lie closely against the curve $\mu = \underline{k}\sigma$. It should be mentioned that this will occur even in the case $1 \leq \varepsilon < 2$, if a 'truncated' density function with $f_z(-\underline{k}-) = 0$ and $f_z(-\underline{k}+) > 0$ prevails. This follows from expressions (2)–(24) in appendix 2.

So far we have only considered the case of distributions bounded to the left. For these distributions the practically relevant part of the (μ, σ) diagram ($\mu/\sigma > -\bar{k}$) is divided into a substitutive and a lexicographic

[25] Cf. section B 1.1.

[26] Cf. the remarks on equation (II D 52).

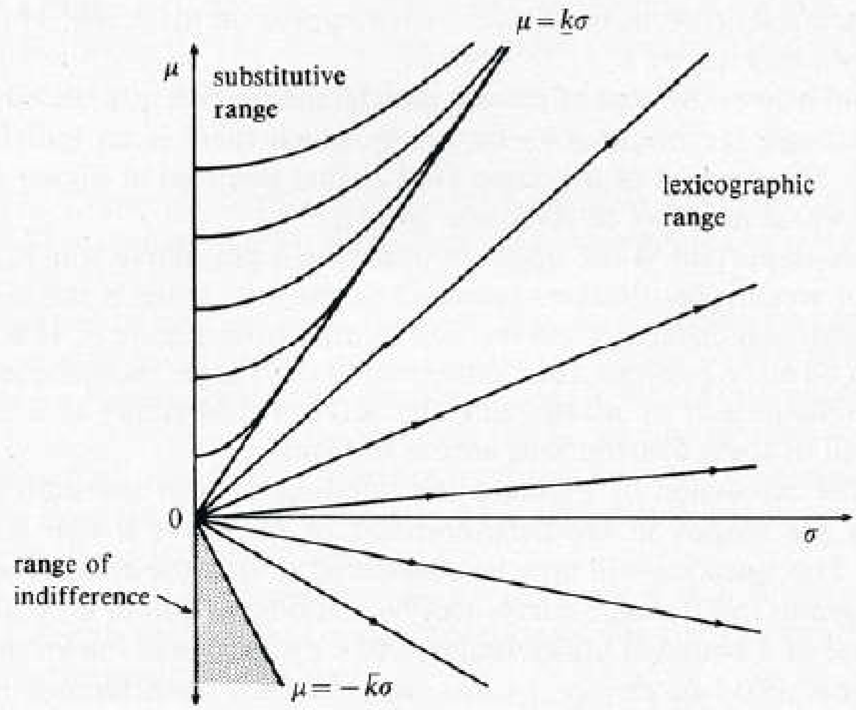


Figure 12

area. The difference when the distributions are, like the normal one, unbounded to the left is immediately clear: the substitutive range disappears completely.

For the case of weak risk aversion ($0 < \varepsilon < 1$) we found that with sufficiently small dispersions risk neutrality prevails so that the optimal decision can be based on expected values alone. This rule is clearly violated in the present case of strong risk aversion ($\varepsilon \geq 1$). In the limit as $\sigma \to 0$ even the slightest increase in standard deviation has to be compensated for by an infinite increase in the expected value. This implication appears highly artificial and suggests that the case $\varepsilon \geq 1$ is not a realistic one. On the other hand it should not be forgotten that not only the net, but also the gross (= balance sheet) distributions of wealth are, in practice, often constrained to the left because there are various forms of limited liability in operation. Even the popular normal distribution is, with respect to its left tail, usually not a good approximation of those gross distributions among which economic decision makers have to choose. Thus there might only be a few occasions where unbounded distributions can be observed.

However, regardless of whether or not the (μ, σ) diagram includes a range of substitutive indifference curves, the implications of the lexicographic range as such are not very plausible. The existence of this range implies that people would be willing to pay an insurance premium of almost their initial wealth to get rid of a liability risk that brings about the possibility of negative gross wealth. Obviously this is rarely the case. People are often unwilling to pay premiums that exceed the expected loss by even a moderate amount; these people, at least, do not have lexicographic preferences. This impression will be reinforced by the multiperiod analysis of chapter IV which shows that *only* the case $\varepsilon < 1$ is compatible with the observation that people become more risk averse as they grow older. Thus there is clear evidence against the preference structure depicted in Figure 12, i.e., against a relative risk aversion greater than or equal to unity. But the evidence is only presumptive. Since we cannot ultimately exclude the possibility that $\varepsilon \geq 1$ will hold for at least some people, the analysis should not be confined to the case $0 < \varepsilon < 1$, however attractive this further reduction in the set of possible preference structures might seem.

An open question in the discussion of Figure 12 is how to choose among distributions with an equal survival probability less than unity. Although the expected utility of all these distributions is $-\infty$, people will not generally be indifferent between them. Indeed, it is possible to find dominance rules that allow an ordering to be made. The distributions considered have the property $\upsilon = \mu + z\sigma$, $E(Z) = 0$, $\sigma(Z) = 1$. This implies that a proportional change in μ and σ which does not affect the probability of survival[27] must be an improvement from the viewpoint of the decision maker. The reason is that (2) and

$$(22) \qquad \lambda\upsilon = \lambda\mu + z\lambda\sigma, \quad \lambda > 1,$$

ensure that each variate z of the standardized random variable Z is associated with a higher variate υ^n of the net wealth distribution if initially $\upsilon^n > 0$, and is associated with the same variate if initially $\upsilon^n = 0$, i.e., if initially gross wealth was zero or negative ($\upsilon \leq 0$). This improvement, which is immediately plausible, follows from the Axioms of Non-Saturation and Independence. According to these axioms, the decision maker is already better off if a single small interval $\underline{z} \leq z \leq \bar{z}$, $\underline{z} < \bar{z}$, can be found where the variates z are associated with higher levels of wealth while elsewhere they bring about given levels of wealth. In Figure 12 this result is reflected by the arrows on the pseudo indifference curves. With strict dominance, a movement along such a curve to the right leads to distributions with a higher evaluation.

The most important aspects of the indifference-curve system in the case $\varepsilon \geq 1$ have now been reported. The results can briefly be summarized.

---

[27] Cf. equations (II B 5) and (II B 6) for $\bar{\upsilon} = 0$.

In the case of strong risk aversion ($\varepsilon \geq 1$), Weber's relativity law in connection with the BLOOS rule implies that, at $\tilde{v}=0$, there is a lexicographic critical level of wealth. Hence maximizing the probability of survival $W(V>\tilde{v})$ is the predominant aim. This aim, however, only has implications for choice if the probability distributions to be evaluated partly extend over the negative half of the wealth axis. If this is not the case, the usual aspects of an evaluation of expected utility remain unaffected.

In the case of linear distribution classes bounded to the left at $\mu - \underline{k}\sigma$, $\underline{k}<\infty$, and to the right at $\mu+\bar{k}\sigma$, $\bar{k}<\infty$, three areas have to be distinguished in the (μ, σ) diagram. An indifference area for $\mu/\sigma \leq -\underline{k}$, an area with rays through the origin as pseudo indifference curves for $-\bar{k}<\mu/\sigma<\underline{k}$, and finally a normal range of substitutive indifference curves for $\mu/\sigma>\underline{k}$. The indifference curves approach the border line between the last two ranges at an angle if $\varepsilon<2$ and if, on the left side of the distributions, density is continuously declining towards zero. The border line is tangent to the indifference curves if $\varepsilon \geq 2$ and/or the probability distribution is truncated at the left-hand side, i.e., if the density jumps to zero. A pseudo indifference curve ranks above another one if it is situated above it. On a pseudo indifference curve, an increasing distance from the origin means that probability distributions with higher evaluations are reached.

In the case of a linear class of unbounded distributions, for example in the case of the class of normal distributions, the whole (μ, σ) diagram is filled with pseudo indifference curves all centering on the origin.

Section C
Arrow's Hypothesis of Increasing Relative and Decreasing Absolute Risk Aversion

ARROW (1965, pp. 28–44; 1970, pp. 90–120) postulates a preference structure that comes, so to speak, half way between the hypotheses of constant absolute and constant relative risk aversion. It implies that an increase in wealth leads to an increase in the intensity of demand for wealth insurance and a decrease in the intensity of demand for insurance of given risk[1].

Crucial to Arrow's argument in favor of the hypothesis of increasing relative risk aversion is his Utility Boundedness Theorem. This theorem requires that, over the positive wealth axis, utility be bounded both from above and below. In deriving his theorem, ARROW (1965, pp. 18–27;

[1] Cf. section A 2.3.2.

1970, pp. 44–89) is attempting to avoid a generalized St. Petersburg Paradox. The St. Petersburg Paradox is an age-old mathematical problem that received its name, which is graphic but not really correct, from the solutions published by BERNOULLI (1738) in St. Petersburg. MENGER (1934) suggested as a resolution to this problem an upper bound to the utility function, thus partly anticipating Arrow's theorem.

Because of

(1) $$\lim_{\upsilon \to 0+} U(\upsilon) = -\infty, \quad \text{if} \quad \varepsilon \geq 1,$$

and

(2) $$\lim_{\upsilon \to +\infty} U(\upsilon) = +\infty, \quad \text{if} \quad \varepsilon \leq 1,$$

the Weber functions (A 34) (cf. also Figure 5) do not meet the postulates of boundedness from below, $\lim_{\upsilon \to 0+} U(\upsilon) > -\infty$, and above, $\lim_{\upsilon \to \infty} U(\upsilon) < +\infty$. Instead, utility functions are required that, for $\upsilon \to 0+$, exhibit a relative risk aversion below and, for $\upsilon \to \infty$, a relative risk aversion above unity, that is, functions that imply *increasing relative risk aversion*[2].

Since Arrow's theorem fundamentally makes the preference hypothesis based on Weber's law doubtful and also because it has been accepted rather uncritically in the literature, it needs to be discussed in more detail. We shall first consider the classical reasoning up to Menger and then move on to Arrow. In connection with Arrow, three problems have to be discussed. First, the question of utility boundedness as such. Second, the question of whether boundedness of utility, if it exists, will have significant implications for the evaluation of risks. And third, the empirical evidence which Arrow thinks he can cite in favor of his hypothesis. Since the third point concerns the optimal structure of asset portfolios its discussion is postponed to chapter V where such questions are considered[3].

---

[2] The corresponding proof can be found in ARROW (1970, pp. 110 f.). A simple formal description of the hypothesis of increasing relative risk aversion can be given by using RUBINSTEIN's (1976) generalized utility function $U(\upsilon) = \ln(\alpha + \upsilon)$ where $\alpha < 0$. This function has the property

$$\varepsilon(\upsilon) = \frac{U''(\upsilon)}{U'(\upsilon)}\,\upsilon = \frac{1}{\alpha/\upsilon + 1} \quad \text{and hence} \quad \frac{\partial \varepsilon(\upsilon)}{\partial \upsilon} > 0.$$

Because of $\lim_{\upsilon \to 0} \ln(\alpha + \upsilon) = \ln \alpha$ the function is bounded from below, but because of $\lim_{\upsilon \to \infty} (\alpha + \upsilon) = \infty$ it is not bounded from above.

[3] Cf. chapter V A 3.3.1.

1. *The St. Petersburg Paradox*

Peter asks Paul how much he would be willing to pay to participate in the following game. A coin is thrown repetitively until 'heads' appears. Then Peter pays Paul an amount of 2^n ducats, where n measures the number of throws. If gambles are evaluated with respect to their expected value, then, because of

$$(3) \qquad E(Y) = \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n 2^n = \sum_{n=1}^{\infty} 1 = \infty,$$

Paul should be willing to pay an infinite stake or at least as much as he owns. That no one behaved this way seemed hard to understand, even paradoxical, from the viewpoint of the classical theory of gambling.

CRAMER (1728) and BERNOULLI (1738), however, believed that they had found an explanation for Paul's behavior in their theory of expected utility[4]. This explanation, they contended, is that

$$(4) \qquad U(a) > \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n U(2^n), \quad U(v) = \begin{cases} \ln v \ \text{(Bernoulli)}, \\ \sqrt{v} \\ \min(v, v^*),\ v^* > 0, \end{cases} \text{(Cramer)},$$

provided Paul's wealth, a, is sufficiently large[5]. Paul in this case would be anxious not to stake his total wealth on the game.

Although no less a person than LAPLACE (1814, p. XV and pp. 439–442) accepted the Cramer-Bernoulli approach, MENGER (1934, esp. p. 468) stated that the utility function $U(v) = \min(v, v^*)$ provides a true solution to the problem but that the functions $\ln v$ and $\sqrt{v}$, like all other functions unbounded from above, do not. In fact, with functions unbounded from above, it is easy to construct a game with

$$(5) \qquad \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n U[\alpha(n)] = \infty.$$

Rather than paying the player an amount 2^n, he simply must get an amount $\alpha(n)$ chosen sufficiently large to ensure that, for all n,

$$(6) \qquad \left(\frac{1}{2}\right)^n U[\alpha(n)] > c > 0, \quad c = \text{const.}$$

[4] Cf. chapter II C 1.2.

[5] In the case of the function $\ln v$, α must be larger than 4 and, in the case of $\sqrt{v}$, larger than $1/(3-2\sqrt{2}) \approx 5.8$. If the units are ducats these are negligible amounts. However, most readers ought to participate in the game if the units are palaces.

Whatever his wealth, Paul should be willing to stake it all for such a game if his utility function is of the type $\ln v$, $\sqrt{v}$, or generally: unbounded from above. Thus there is a new 'paradox'. A genuine solution, Menger argued, can only be found when there is an upper bound to the utility function as, for example, in Cramer's second function. In this case (6) cannot be satisfied.

Thus it seems that only an upper bound to the utility function ensures that the implications of the expected-utility rule are compatible with the true bahavior of man. However, other solutions to the St. Petersburg Paradox have been offered, showing that this conclusion is too hasty[6].

CHIPMAN (1960, p. 221) tried to explain the low level of the stake (P) for the St. Petersburg game by a lexicographic critical level of wealth $\tilde{v} > 0$ which requires $a - P + \alpha(n) > \tilde{v}$. FURLAN (1946) believed the solution is to discount the prizes since playing the game takes time. MENGER (1934, pp. 471 f.) also considered the possibility of explaining the paradox by the fact that people tend to neglect small probabilities. SENETTI (1976) argues that, for the St. Petersburg game, the coefficient of variation is infinite, so that, according to the usual shapes of indifference curves in a (μ, σ) diagram, a participation is not attractive. Unfortunately he does not show the relationship with the von Neumann-Morgenstern function that he implicitly assumes. This, however, would be necessary since the very unusual shape of the probability distribution of prizes in the St. Petersburg game suggests very unusual shapes of the indifference curves.

A simple solution was proposed by Bernoulli's German translator Pringsheim (BERNOULLI (1738, German edition 1896, fn. 10, pp. 46–52)) but, as TODHUNTER (1865, p. 222) and KEYNES (1921, p. 317) indicate, it really dates back to Poisson, Concordet, and Cournot. According to this solution Peter simply offered too much. If Paul is smart, he knows that, at around about $k = 50$ throws, the whole wealth of the world would not cover the prize promised by the St. Petersburg game. Since Peter would go broke even sooner, the expected prize of the game is definitely finite, that is, less than $k+1$ ducats[7,8]. Hence Paul would be

---

[6] An extensive overview of the literature is given by SAMUELSON (1977).

[7] If 2^k is the maximum prize Peter can pay, then the expected prize as calculated by Paul is

$$E(Y) = \sum_{n=1}^{k} (1/2)^n 2^n + \left[1 - \sum_{n=1}^{k} (1/2)^n\right] 2^k,$$

or, after some basic transformations, $E(Y) = k + 1$.

[8] Another reason for the finiteness of the effective prize is given by BRITO (1975). On the basis of Becker's theory of consumption he formally 'proves' that finiteness is required because more time than the gambler has available may be necessary in order to consume the prize.

wise to think twice before risking his wealth. This argument can hardly be refuted. BERTRAND's (1907, p. 61) objection, that it is possible to have the expected gain approaching infinity by reducing to zero the unit in which the prize is paid, is not to the point, for in this way it is not possible to induce Paul to stake his total wealth. Of course, the expected prize approaches infinity when it is measured in the reduced unit, but, an aspect that is often overlooked, when it is measured in the initial unit (ducats) it approaches zero[9]. This would provide an even stronger reason for Paul being unwilling to give away his wealth[10,11].

Thus it can be concluded that there is no need to solve the classical version of the St. Petersburg Paradox by introducing the expected-utility concept let alone by adding the assumption that utility is bounded from above. The true solution is to be found in the limitation of the prizes.

2. *The Utility Boundedness Theorem*

In a somewhat abbreviated form, ARROW's (1970, pp. 63–69) reasoning in favor of the boundedness of utility runs as follows. Let e_1 be a probability distribution over the strictly positive half of the wealth axis that has a finite number of variates. Then, with a utility function that is defined for all strictly positive values of its argument, the utility of a single variate is finite and so is the expected utility. Moreover, let e_2 be a probability distribution of Menger's type (6) that has an infinite number of variates and that brings about an expected utility of infinity. In addition, let there be a third distribution, e_3, that is also of the type (6) but offers prizes $\bar{\alpha}(n)$, where $\bar{\alpha}(n) > \alpha(n)$ for all n and where $\alpha(n)$ denotes the prizes of distribution e_2. The preference ordering over the three distributions will then be $e_1 < e_2 < e_3$, where $e_2 < e_3$ follows from a dominance axiom postulated by ARROW (1970, p. 50) or from the Axiom of Strong Independence used in this book. Finally, Arrow defines a further distribution e_4 that, together with e_2 and e_3, is represented in the following table. This distribution has the property that, starting with the

---

[9] Suppose the unit is reduced to $1/(2^x)$ ducats. Then, in terms of the new unit (N), Peter's wealth is 2^{k+x} so that, according to the formula of footnote 7, the expected prize in terms of new units is $E(Y_N) = k + x + 1$. Expressed in ducats (d), this is equivalent to an expected prize of $E(Y_d) = (k + x + 1)/2^x$. Hence $\lim_{x \to \infty} E(Y_d) = 0$.

[10] Cf. e.g., KEYNES (1921, p. 317) and GOTTINGER (1971/72, p. 494) who seem to accept Bertrand's objection.

[11] SAMUELSON (1960) shows that, measured in current units, the maximum stake of an expected-utility maximizer approaches infinity as the unit reduces to zero. This result, too, is not to the point.

jth throw, only the arbitrarily choosable amount β, $\beta \geq 0$, is paid out, while before this throw the prizes equal those of distribution e_3, i.e., $\bar{\alpha}(n)$.

throw	1	2	...	$j-1$	j	$j+1$	...
probability	$\left(\frac{1}{2}\right)^1$	$\left(\frac{1}{2}\right)^2$	...	$\left(\frac{1}{2}\right)^{j-1}$	$\left(\frac{1}{2}\right)^j$	$\left(\frac{1}{2}\right)^{j+1}$	...
e_2	$\alpha(1)$	$\alpha(2)$	...	$\alpha(j-1)$	$\alpha(j)$	$\alpha(j+1)$	...
e_3	$\bar{\alpha}(1)$	$\bar{\alpha}(2)$	...	$\bar{\alpha}(j-1)$	$\bar{\alpha}(j)$	$\bar{\alpha}(j+1)$	...
e_4	$\bar{\alpha}(1)$	$\bar{\alpha}(2)$	...	$\bar{\alpha}(j-1)$	β	β	...

Obviously distribution e_4 has a finite number, j, of variates and hence, like distribution e_1, must be worse than e_2: $e_2 \succ e_4$, On the other hand, it is to be expected that, by choosing j sufficiently large, the evaluation of e_4 will approach that of e_3 as closely as we wish, for, with an increase in j, the probability that e_4 will bring about an outcome different from e_3 approaches zero. Thus the assumption $e_3 \succ e_2$ suggests that ultimately, with a very large j, we find $e_4 \succ e_2$. This, however, would be a contradiction which Arrow believes can be avoided only if an upper bound on the utility curve is postulated since such a bound ensures that e_2 and e_3 bring about a finite level of expected utility.

It was seen from the discussion of the classical version of the St. Petersburg Paradox that the contradiction constructed by Arrow cannot happen in the real world since distributions e_2 and e_3 with prizes approaching infinity do not exist. If Arrow, nevertheless, insists on determining the shape of the utility curve for wealth levels impossible in the real world then he can change the unbounded Weber functions $U(v) = \ln v$ and $U(v) = (1-\varepsilon)v^{1-\varepsilon}$, $\varepsilon < 1$, for wealth levels above the value of the whole wealth of the world. That should make him happy and us too.

Despite this, Arrow's argument cannot be labelled as completely irrelevant for practical decision problems under uncertainty, for it can be used to legitimate a lower bound on utility in the same way as it was used to legitimate an upper bound. If the utility function is unbounded from below we may, similarly to (6), construct a *vanishing* sequence of variates $\alpha(n)$ of a wealth distribution such that

$$(7) \qquad \left(\frac{1}{2}\right)^n U[\alpha(n)] < c < 0, \quad c = \text{const.}, \quad \alpha(n) > 0.$$

In this case, there is a distribution e_2 with an expected utility of $-\infty$, which is smaller than the expected utility of any other distribution e_1

that has a finite number of strictly positive wealth variates so that $e_1 \succ e_2$. If then, analogously to the previous procedure, distributions e_3 with $\bar{\alpha}(n) < \alpha(n)$ and e_4 are constructed, a contradiction arises. On the one hand $e_4 \succ e_2$, since e_4 has a finite number of variates, on the other hand $e_4 \prec e_2$. The latter is so because, with an increase in j, e_4 can be made as 'similar' as we like to e_3 and e_3 is definitely worse than e_2. The contradiction can be removed if the utility function is bounded at $v \to 0+$ since, in this case, e_2 and e_3 both bring about a finite level of expected utility.

It was argued that Arrow's postulate of an upper bound to utility is meaningless since in the real world wealth is bounded from above. The assumption of a lower bound cannot be discredited on similar grounds because wealth levels down to zero can easily be realized[12]. Thus the decisive question is whether it makes sense to assume that, in the preference ordering of the decision maker, it is indeed possible to have e_4 approaching e_3 as closely as we wish by choosing j sufficiently large.

This is exactly what is postulated in ARROW's (1970, pp. 48 f.) *Monotone Continuity Axiom*. This axiom is similar to the *Archimedes Axiom* introduced above[13], but not identical. While the Monotone Continuity Axiom refers to a comparison of two probability distributions one of which has an infinite number of variates, the Archimedes Axiom refers to a comparison between a non-random level of wealth and a binary distribution requiring that there be some probability w, $0 < w < 1$, such that $U(\bar{v}) = wU(v_1) + (1-w)U(v_2)$, $v_1 < \bar{v} < v_2$. To understand clearly the difference between the two axioms, consider the following two possibilities.

1. The axioms are postulated for wealth levels greater than or equal to zero. In this case, using Arrow's reasoning, we can choose some arbitrary $\beta \geq 0$ and then we find that utility has to be bounded for $v \to 0$. The same implication holds for the Archimedes Axiom, for otherwise, with $v_1 = 0$, an indifference probability in the open unit interval would not exist.

2. The axioms are postulated for strictly positive wealth levels. Obviously this makes no difference to Arrow's argument since it holds equally well if β is limited to being strictly positive. It does make a substantial difference, however, when the Archimedes Axiom is assumed. If $v_1 > 0$, there exists an indifference probability in the closed unit interval even though $U(v)$ is not bounded for $v \to 0$. In particular, such

[12] ARROW's (1974a) rejoinder to an objection of RYAN (1974) points in this direction. Arrow states that the existence of the mathematical expectation of the probability distribution to be evaluated ensures the existence of expected utility if the utility function is monotonically increasing, concave, and bounded from below.

[13] Cf. chapter II C 2.1.

an indifference probability exists for the unbounded Weber functions $U(v) = \ln v$ and $U(v) = (1-\varepsilon)v^{(1-\varepsilon)}$, $\varepsilon > 1$, since $U(v_1)$, $U(\bar{v})$, and $U(v_2)$ are all well-defined and finite.

Thus it turns out to be a peculiarity of the Monotone Continuity Axiom that it implies boundedness of utility even if the wealth variates to which the axiom refers are limited to being strictly positive. The Archimedes Axiom implies a boundedness of utility only when it is assumed to hold for all wealth levels including zero. The question of whether or not the Archimedes Axiom should include zero levels of wealth therefore is the question of whether or not a lexicographic critical wealth level at $v = \bar{v} = 0$ should be excluded. We have seen that there are good empirical reasons for regarding lexicographic preferences as atypical[14]. Apart from this, however, there do not appear to be any compelling reasons for totally excluding these preferences. Unless we want to extend our axiom system by the additional axiom that St. Petersburg gambles bring about a finite level of expected utility, there is no need for utility to be bounded.

Occasionally Arrow's findings have been superficially interpreted as implying that *the* expected-utility axioms require a boundedness of utility. This interpretation is fallacious. The axiom system used in this book does imply the expected-utility rule, but it by no means excludes utility functions unbounded from above or from below.

3. *The Missing Behavioral Implications of the Utility Boundedness Theorem*

Regardless of the fact that our axioms do not imply a boundedness of utility, we want to accept Arrow's postulates for the moment and to look into their behavioral implications for the choice between probability distributions which are less awkward than those of the generalized St. Petersburg type. Arrow (1970, p. 98) himself sees the two following implications: '(1) it is broadly permissible to assume that relative risk aversion increases with wealth, though theory does not exclude some fluctuations; (2) if, for simplicity, we wish to assume a constant relative risk aversion, then the appropriate value is one.' These implications are obviously formulated in rather cautious language. Nevertheless, the formulation suggests far more economic content than is really implied by the boundedness of utility. In order to show this, let us try to modify the Weber functions in such a way that, in the limit as $v \to 0$ or $v \to \infty$, they satisfy Arrow's postulates while the evaluation of the probability

[14] Cf. section B 2 and the discussion of the lexicographic criteria in chapter II B.

distribution under consideration is altered only to a negligible extent. If Arrow's Utility Boundedness Theorem has empirically relevant implications then this attempt must fail, at least for the Weber function $U(v) = (1-\varepsilon)v^{1-\varepsilon}$, $\varepsilon \neq 1$, excluded by Arrow.

Assume first that $\varepsilon < 1$ and define, in line with what has been suggested in the previous section, the modified function

$$(8) \qquad U_m(v) \equiv \begin{Bmatrix} U(v), & v < \bar{v} \\ U(\bar{v}), & v \geq \bar{v} \end{Bmatrix}, \quad 0 < \varepsilon < 1,$$

$$\bar{v} \equiv \text{all the world's wealth.}$$

Then, for all economic distributions possible in the real world the upper bound to utility has no behavioral implications whatsoever. There is no reason why it should be 'broadly permissible' to assume increasing relative risk aversion for the relevant range, nor is it clear why the appropriate function is logarithmic if 'for simplicity' we want to assume that relative risk aversion is constant.

At first glance the matter seems to be different if the Weber functions for $\varepsilon \geq 1$, which are unbounded from below, are modified so that

$$(9) \qquad U_m(v) \equiv \begin{Bmatrix} U(v), & v > \underline{v} \\ U(\underline{v}), & v \leq \underline{v} \end{Bmatrix}, \quad \varepsilon \geq 1, \quad \underline{v} > 0.$$

Of course this modification has no implications for probability distributions that are entirely above $\underline{v}$. But implications are to be expected in the case of distributions that partly extend below $\underline{v}$. In particular, it might be expected that gross distributions that incorporate the possibility of negative wealth are drastically affected since

$$U_m(0) = U_m(\underline{v}) = U(\underline{v}) > -\infty.$$

The expectation of drastic implications is wrong however. This will be proved by showing that, by choosing $\underline{v}$ sufficiently small, it is possible to define the function $U_m(v)$ in such a way that it generates the same preference ordering over two gross distributions V_1 and V_2 as the one implied by $U(v)$

(1) if both distributions extend only over the strictly positive half of the wealth axis so that for each of them we have $\mu - \underline{k}\sigma > 0$ where $\underline{k}$ is the lower boundary of the standardized random variable characterizing the distribution in question;
(2) if the distributions indicate different survival probabilities

$W(v > 0)$ (at least one of the distributions then extends partly over the negative half of the wealth axis);

(3) if both distributions belong to the same linear class and yield the same survival probability while their standard deviations differ.

Referring to (1): To show this, is a trivial task. We simply have to choose $\underline{v}$ sufficiently small so that, over the range of wealth covered by the two distributions, $U(v) = U_m(v)$. This is shown in Figure 13 for the example of a linear distribution class. The postulate $\mu - \underline{k}\sigma > \underline{v}$ gives a lower boundary line for the range where the indifference curves generated by $U_m(v)$ have the same shape as those generated by $U(v)$. If $\underline{v}$ is reduced, then this boundary line can be made to approach the former boundary $\mu = \underline{k}\sigma$ to the area of lexicographic indifference curves as closely as we wish.

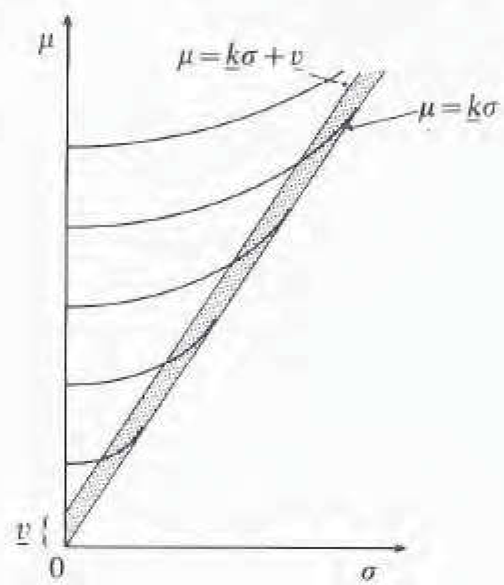


Figure 13

Referring to (2): It was argued above that a combination of the Bloos rule and the Weber functions unbounded from below implies that maximizing the probability of survival is the predominant aim[15]. Suppose the contention (2) is correct. Then, comparing two different probability distributions V_1 and V_2, we must find that, when $\underline{v}$ approaches zero, at some stage the distribution with the higher survival probability definitely brings about the higher level of expected utility, utility being given by the modified function $U_m(v)$. Let us thus calculate

[15] Cf. section B 2.

$$(10)\qquad D \equiv \int_{-\infty}^{+\infty} f_1(v)U_m(v)\,dv - \int_{-\infty}^{+\infty} f_2(v)U_m(v)\,dv$$
$$\equiv U(\underline{v}) \int_{-\infty}^{\underline{v}} f_1(v)\,dv + \int_{\underline{v}}^{+\infty} f_1(v)U(v)\,dv$$
$$- U(\underline{v}) \int_{-\infty}^{\underline{v}} f_2(v)\,dv - \int_{\underline{v}}^{+\infty} f_2(v)U(v)\,dv,$$

which is the difference between the expected utilities of the two distributions with $f_1(v)$ and $f_2(v)$ as the corresponding density functions. We then can discover what the relationship

$$(11)\qquad V_1 \left\{\begin{smallmatrix}\succ\\ \sim\\ \prec\end{smallmatrix}\right\} V_2 \Leftrightarrow D \left\{\begin{smallmatrix}>\\ =\\ <\end{smallmatrix}\right\} 0$$

has in common with the relationship

$$(12)\qquad V_1 \left\{\begin{smallmatrix}\succ\\ \sim\\ \prec\end{smallmatrix}\right\} V_2 \Leftrightarrow W(V_1>0) = \int_0^{\infty} f_1(v)\,dv \left\{\begin{smallmatrix}>\\ =\\ <\end{smallmatrix}\right\} W(V_2>0) = \int_0^{\infty} f_2(v)\,dv.$$

Assume that $\underline{v}$ is chosen sufficiently small so as to ensure that $U(\underline{v})<0$. Then, dividing (10) by $U(\underline{v})$, rearranging terms, separating the integrals $\int_{\underline{v}}^{\infty}\ldots dv$ into the sum $\int_{\underline{v}}^{\underline{v}'}\ldots dv + \int_{\underline{v}'}^{\infty}\ldots dv$, $0<\underline{v}<\underline{v}'$, $U(\underline{v})<U(\underline{v}')<0$, and taking into consideration the fact that $\int_{-\infty}^{\underline{v}} f_i(v)\,dv = 1 - \int_{\underline{v}}^{\infty} f_i(v)\,dv$, $i=1,2$, we have[16]

$$(13)\qquad \operatorname{sgn} D = \operatorname{sgn}\left[\int_{\underline{v}}^{\infty} f_1(v) - f_2(v)\,dv - \int_{\underline{v}}^{\underline{v}'} [f_1(v)-f_2(v)]\frac{U(v)}{U(\underline{v})}\,dv - \int_{\underline{v}'}^{\infty} [f_1(v)-f_2(v)]\frac{U(v)}{U(\underline{v})}\,dv\right].$$

With $\underline{v}\to 0$, given $\underline{v}'$, the last integral vanishes because

$$(14)\qquad \lim_{\underline{v}\to 0}\frac{U(v)}{U(\underline{v})} = \left\{\begin{array}{ll} \lim\limits_{\underline{v}\to 0}\dfrac{\ln v}{\ln \underline{v}} = 0, & \varepsilon = 1 \\ \lim\limits_{\underline{v}\to 0}\dfrac{(1-\varepsilon)v^{1-\varepsilon}}{(1-\varepsilon)\underline{v}^{1-\varepsilon}} = 0, & \varepsilon > 1 \end{array}\right\}\ \forall v \ge \underline{v}' > \underline{v}.$$

[16] Cf. footnote 36 in chapter II D.

The influence of the second integral can also be made arbitrarily small. It is true that when $\underline{v}$ approaches zero its change in value is ambiguous. However, there is an upper bound to this integral which, by a suitable initial choice of $\underline{v}$ and $\underline{v}'$, can be made as small as we wish without changing the validity of (14). Because of $U(v)/U(\underline{v}) \leq 1$ this upper bound is given by

$$(15) \qquad \left| \lim_{\underline{v} \to 0} \int_{\underline{v}}^{\underline{v}'} [f_1(v) - f_2(v)] \frac{U(v)}{U(\underline{v})} dv \right| \leq \left| \lim_{\underline{v} \to 0} \int_{\underline{v}}^{\underline{v}'} [f_1(v) - f_2(v)]\, dv \right| .$$

Thus, only the first integral remains. With $\underline{v} \to 0$, its value approaches the difference in the survival probabilities of the two distributions. Thus, overall, we have

$$(16) \qquad \operatorname{sgn}\left[\lim_{\underline{v}' \to 0} \lim_{\underline{v} \to 0} D \right] = \operatorname{sgn}\left[\int_0^{\infty} f_1(v)\, dv - \int_0^{\infty} f_2(v)\, dv \right],$$

which gives the relationship between (11) and (12) that we sought. The result is worth noting for it indicates that, by the use of the bounded function $U_m(v)$, it is possible to approximate the predominance of the survival probability, as implied by the unbouned Weber functions, as closely as we wish. In the case of linear distribution classes, to which this result is not, however, limited, this leads to the clear interpretation that, in the range $\mu/\sigma < \underline{k}$ of the (μ, σ) diagram, a decrease in $\underline{v}$ makes the real indifference curves approach more and more closely the pseudo indifference curves described above[17].

Referring to (3): For the unmodified Weber functions it was shown that, if there are two distributions that exhibit the same survival probability and belong to the same linear class, the distribution with the higher standard deviation is to be preferred[18]. The remaining question, therefore, is whether this result continues to hold when $U_m(v)$ is used. The answer can easily be given if, within (10), the first distribution is extended by multiplying the single variates with a factor of proportionality λ, $\lambda > 1$. This multiplication does not affect the survival probability[19] but, because of

[17] This can also be proved by using equation (A 48).

[18] Cf. section B 2.

[19] Cf. equations (II B 5) and (II B 6) for $\bar{v} = 0$.

$$D = U(\underline{v}) \int_{-\infty}^{\underline{v}/\lambda} f_1(v)\,dv + \int_{\underline{v}/\lambda}^{\infty} f_1(v)U(\lambda v)\,dv - U(\underline{v}) \int_{-\infty}^{\underline{v}} f_2(v)\,dv - \int_{\underline{v}}^{\infty} f_2(v)U(v)\,dv \tag{17}$$

where

$$\frac{dD}{d\lambda} = \int_{\underline{v}/\lambda}^{\infty} f_1(v)U'(\lambda v)\,v\,dv > 0, \tag{18}$$

the difference in expected utilities rises irrespective of $\underline{v}$. If we assume that initially $D = 0$ and note that $\sigma(\lambda V) = \lambda\sigma(V)$ then this result confirms contention (3). In a certain sense it seems to be the opposite of the one derived under (2). How is it possible on the one hand for the indifference curves to approximate rays through the origin in the (μ, σ) diagram while, on the other hand, a movement along these rays away from the origin leads to higher-ranking indifference curves? The answer is simply that, while with $\underline{v} \to 0$ the slopes of the indifference curves approach those of the corresponding rays through the origin, they never exactly coincide when $\underline{v} > 0$. In the (μ, σ) diagram, this property can only be represented indirectly, for example, by attaching outward pointing arrows to the indifference curves, as we did in the case of pseudo indifference curves. Thus we can just as well maintain the indifference-curve system depicted in Figure 12 which was derived from the unbounded Weber functions.

So the verdict on Arrow's hypothesis of increasing relative risk aversion is: even if the assumptions that imply a boundedness of utility are accepted, there are practically no behavioral implications for the evaluation of probability distributions. Although all types of Weber functions are unbounded at least in one direction, they can easily be modified so that utility is bounded but the evaluation of risk projects either does not change at all or, if it does, the change can be made as small as we wish. This means that, with respect to the results of the two preceding sections also, we should join with Samuelson (1969, p. 243) when he says: 'Since I do not believe that Karl Menger paradoxes of the generalized St. Petersburg type hold any terrors for the economist, I have no particular interest in boundedness of utility ...'

Appendix 1 to Chapter III

Note first that a density function $f_{w_1}(w)$ brings about a higher mathematical expectation $E(W)$ than another function $f_{w_2}(w)$ if there is a w^* such that

$$\text{(1)}\qquad \begin{aligned} &f_{w_1}(w) > f_{w_2}(w), \quad \text{if} \quad w > w^*,\\ &\text{and}\\ &f_{w_1}(w) < f_{w_2}(w), \quad \text{if} \quad w < w^* \end{aligned}$$

for all w where $\min(f_{w_1}, f_{w_2}) > 0$. It is assumed that the mathematical expectations for the two density functions are finite.

The problem is to calculate the sign of the derivative $d\Psi/d\varepsilon$ where

$$\text{(2)}\qquad \psi = -\frac{\int\limits_{-\mu/\sigma}^{\infty} z f_z(z)\left(\frac{\mu}{\sigma}+z\right)^{-\varepsilon} dz}{\int\limits_{-\mu/\sigma}^{\infty} f_z(z)\left(\frac{\mu}{\sigma}+z\right)^{-\varepsilon} dz}.$$

To prepare for this task, choose a number $x > 0$ such that, given another number Δ, $0 < \Delta < \infty$,

$$\text{(3)}\qquad \int\limits_{-\mu/\varrho}^{\infty} f_z(z)\left(\frac{z+\mu/\sigma}{x}\right)^{-\varepsilon-\Delta} dz = \int\limits_{-\mu/\sigma}^{\infty} f_z(z)\left(\frac{\mu}{\sigma}+z\right)^{-\varepsilon} dz.$$

Such a number can always be found, provided that both integrals in (3) are finite, since the left-hand integral approaches zero or infinity as x goes to zero or infinity. Since $\Delta > 0$ and $\varepsilon > 0$ the right-hand integral is clearly finite if the left-hand integral is. Two conditions that are sufficient for a finiteness of the left-hand integral can be obtained from appendices 2 and 4 below. The first is

$$\text{i)}\qquad f_z\left(-\frac{\mu}{\sigma}+\right) \geq 0, \quad \varepsilon+\Delta < 1.$$

It follows from expressions (4) and (20) in appendix 2. (Substitute $\mu/\sigma \equiv y$, $z \equiv w$, $\varepsilon+\Delta \equiv \Theta$.) The second is

$$\text{ii)}\qquad f_z\left(-\frac{\mu}{\sigma}+\right) = 0, \quad \varepsilon+\Delta < 2.$$

This condition follows from expression (13) in appendix 4. (Substitute $z \equiv w$, $\mu/\sigma \equiv y$, $\varepsilon+\Delta \equiv \Theta$.) Condition i) is relevant for expressions

(III A 53) and (III B 18) of the text. Condition ii) will be referred to in appendix 3.

Consider now the following definition

$$(4) \qquad \psi_\Delta \equiv -\frac{\int_{-\mu/\sigma}^{\infty} z f_z(z)\left(\frac{z+\mu/\sigma}{x}\right)^{-\varepsilon-\Delta} dz}{\int_{-\mu/\sigma}^{\infty} f_z(z)\left(\frac{z+\mu/\sigma}{x}\right)^{-\varepsilon-\Delta} dz}.$$

If we reduce the quotient on the right-hand side by $x^{\varepsilon+\Delta}$ then we get a formula very similar to (2), the only difference being that ε is replaced by $\varepsilon+\Delta$. Hence we can conclude that

$$(5) \qquad \frac{d\psi}{d\varepsilon}\left\{\begin{smallmatrix}\geq\\=\\\leq\end{smallmatrix}\right\} 0 \Leftrightarrow \psi_\Delta \left\{\begin{smallmatrix}\geq\\=\\\leq\end{smallmatrix}\right\} \psi.$$

A comparison between ψ_Δ and ψ is not difficult. Since, by construction, the denominators of both expressions are equal, we only have to consider the numerators. From the self-evident assumption $\int_{-\mu/\sigma}^{\infty} z f_z(z)\,dz < \infty$ and the finiteness of the denominator, it follows that the numerator is finite, too. Identifying $f_z(z)(.)^{-\varepsilon-\delta}/\int_{-\mu/\sigma}^{\infty} f_z(z)(.)^{-\varepsilon-\delta} dz$ $(\delta=0,\Delta)$ with the density functions f_{w_1} and f_{w_2} respectively, then, from (1), we obviously have $-\psi > -\psi_\Delta$, i.e.,

$$(6) \qquad \psi_\Delta > \psi$$

if there exists a value z^* such that

$$f_z(z)\left(z+\frac{\mu}{\sigma}\right)^{-\varepsilon}\left\{\begin{smallmatrix}\geq\\=\\\leq\end{smallmatrix}\right\} f_z(z)\left(\frac{z+\mu/\sigma}{x}\right)^{-\varepsilon-\Delta}$$

$$\Leftrightarrow z\left\{\begin{smallmatrix}\geq\\=\\\leq\end{smallmatrix}\right\} z^* \quad \text{and} \quad f_z(z)>0,\ z\geq-\mu/\sigma.$$

Because of $\Delta>0$ and by the definition of x, this is clearly the case. Hence $d\psi/d\varepsilon>0$, provided conditions i) and/or ii) are met.

Appendix 2 to Chapter III

We attempt to find out how the value of the quotient

$$(1) \qquad A = -\frac{\int_{-y}^{\infty} w f_w(w)(y+w)^{-\Theta}\,dw}{\int_{-y}^{\infty} f_w(w)(y+w)^{-\Theta}\,dw}$$

is related to y. For this purpose we first consider the quotient

$$B = -\frac{\int_{-y}^{\infty} w f_w(w)(x+w)^{-\Theta}\, dw}{\int_{-y}^{\infty} f_w(w)(x+w)^{-\Theta}\, dw}, \quad x > y, \tag{2}$$

and identify its limit with the quotient whose value is to be calculated:

$$\lim_{x \to y+} B = A. \tag{3}$$

The variables x, y, and Θ are strictly positive real numbers and $f_w(w)$ is a probability density function for the random variable W. The variable y is defined such that $w = -y$ is the highest lower bound on w. It is assumed for the time being that the density function is 'truncated' at $w - y$, i.e., that[1]

$$\lim_{w \to -y-} f_w(w) \equiv f_w(-y-) = 0 \quad \text{and} \quad \lim_{w \to -y+} f_w(w) \equiv f_w(-y+) > 0. \tag{4}$$

In some arbitrarily small region $-y < w < -y + \Delta$, $\Delta > 0$, $f_w(w)$ is to be continuous, strictly positive, and either monotonically increasing or decreasing. Moreover, it is assumed that

$$\int_{0}^{\infty} w f_w(w)\, dw < \infty. \tag{5}$$

We now write (2) as

$$B = -[\alpha(1-\gamma) + \beta\gamma] \tag{6}$$

with

$$\alpha \equiv \frac{\int_{-y}^{-y+\Delta} w f_w(w)(x+w)^{-\Theta}\, dw}{\int_{-y}^{-y+\Delta} f_w(w)(x+w)^{-\Theta}\, dw}, \tag{7}$$

$$\beta \equiv \frac{\int_{-y+\Delta}^{\infty} w f_w(w)(x+w)^{-\Theta}\, dw}{\int_{-y+\Delta}^{\infty} f_w(w)(x+w)^{-\Theta}\, dw}, \tag{8}$$

[1] $f_w(-y+)$ is the right-hand side density at $w = -y$.

and

$$(9) \qquad \gamma \equiv \frac{\int_{-y+\Delta}^{\infty} f_w(w)(x+w)^{-\Theta}\,dw}{\int_{-y}^{\infty} f_w(w)(x+w)^{-\Theta}\,dw},$$

where Δ is chosen such that $0 < \gamma < 1$.

With α and β, conditional arithmetic means of W are defined. Since α only covers the range $-y \leq w \leq -y + \Delta$ and β the range $w \geq -y + \Delta$, we have

$$(10) \qquad -y < \alpha < -y + \Delta < \beta.$$

Note that, before the limit $x \to y+$ is taken, $x + w > 0\ \forall w$ and hence $0 < (x+w)^{-\Theta} < \infty$. Consider now the limit

$$(11) \qquad \lim_{x\to y+} B = -\left[\lim_{x\to y+} \alpha\left(1 - \lim_{x\to y+} \gamma\right) + \lim_{x\to y+} \beta \lim_{x\to y+} \gamma\right].$$

Concerning α, it is sufficient to state that it may reach its boundaries but cannot exceed them:

$$(12) \qquad -y \leq \lim_{x\to y+} \alpha \leq -y + \Delta.$$

Concerning β, more precise information is available. The value of β is determined by variates of w for which $w \geq -y + \Delta$ and hence $x + w > x - y + \Delta > 0$. This implies $(x+w)^{-\Theta} < \infty$ even if, with $x = y$, the limit has been taken. Hence

$$(13) \qquad \lim_{x\to y+} \beta > -y + \Delta.$$

We now consider the limit of γ. Note first that, for the reason just discussed in connection with β, we have $(x+w)^{-\Theta} < \infty$. Thus the numerator is strictly positive and finite. The denominator, call it N, is strictly positive, but it is not clear *a priori* whether it is finite. Obviously, for the value of the denominator, the boundaries

$$(14) \qquad N_1 + N_3 \leq N \leq N_2 + N_3, \quad N_1 \leq N_2,$$

or

$$(15) \qquad N_1 + N_3 \geq N \geq N_2 + N_3, \quad N_1 \geq N_2,$$

will prevail where

$$N_1 \equiv f_w(-y+) \int_{-y}^{-y+\Delta} (x+w)^{-\Theta}\, dw, \tag{16}$$

$$N_2 \equiv f_w(-y+\Delta) \int_{-y}^{-y+\Delta} (x+w)^{-\Theta}\, dw, \tag{17}$$

and

$$N_3 \equiv \int_{-y+\Delta}^{\infty} f_w(w)(x+w)^{-\Theta}\, dw. \tag{18}$$

When taking the limit, N_3 certainly stays finite, since $x+w+\Delta>0$ even when $x=y$. The matter is different with N_1 and N_2. Because of

$$\int_{-y}^{-y+\Delta} (x+w)^{-\Theta} = \begin{cases} \ln(x-y+\Delta)-\ln(x-y),\ \Theta=1, \\ (1-\Theta)[x-y+\Delta)^{1-\Theta}-(x-y)^{1-\Theta}],\ \Theta\neq 1, \end{cases} \tag{19}$$

the limits of these two integrals satisfy

$$0<\lim_{x\to y+} N_{1,2} \begin{cases} <\infty,\ \Theta<1, \\ =\infty,\ \Theta\geq 1. \end{cases} \tag{20}$$

From this expression, the finiteness of the numerator, the finiteness of N_3, and the fact that the numerator is clearly smaller than the denominator it follows for the limit of γ that

$$1>\lim_{x\to y+} \gamma \begin{cases} >0,\ \Theta<1, \\ =0,\ \Theta\geq 1. \end{cases} \tag{21}$$

With the aid of the information contained in (12), (13), and (21) we are now able to address the problem of determining the value of A. If $\Theta<1$, the limit of B is a negative linear combination of $\lim\alpha$ and $\lim\beta$ with weights strictly between zero and one. Since $\lim\beta>-y$ will hold even when $\Delta\to 0$, we have $-\lim B>-y$ and, because $\lim_{x\to y+} B=A$,

$$A<y,\quad \Theta<1;\ f_w(-y+)>0. \tag{22}$$

If, on the other hand, $\Theta\geq 1$ then β gets the weight zero and the boundaries of α also apply to $-\beta$. Thus

$$y-\Delta\leq A\leq y,\quad \Theta\geq 1. \tag{23}$$

Now, the above reasoning was not confined to a particular value of Δ. We can therefore choose Δ as small as we wish and must hence conclude that

$$(24)\qquad A=y,\quad \Theta\geq 1;\quad f_w(-y+)>0.$$

Having achieved this result for the case of 'truncated' density, we next approach the case where at $w=-y$ the density vanishes continuously. Instead of (4) it is therefore assumed that

$$(25)\qquad \lim_{w\to -y-} f_w(w)\equiv f_w(-y-)=\lim_{w\to -y+} f_w(w)\equiv f_w(-y+)=0,$$
$$0<f'_w(0+)<\infty.$$

The only part, which has to be changed in the above reasoning, is the one in which the size of the denominator N of the quotient on the right-hand side of equation (9) is calculated. In appendix 4 it is shown (by the reasoning up to expression (14)) that for

$$(26)\qquad \lim_{x\to y+} N=\int_{-y}^{\infty} f_w(w)(y+w)^{-\Theta}\,dw$$

it holds that

$$(27)\qquad \lim_{x\to y+} N\begin{Bmatrix}<\\=\end{Bmatrix}\infty \Leftrightarrow 0<\Theta\begin{Bmatrix}<\\\geq\end{Bmatrix}2.$$

Thus, completely analogously to the above reasoning, instead of (22) and (24), we have

$$(28)\qquad A<y,\quad \Theta<2,\quad f_w(-y)=0,$$

and

$$(29)\qquad A=y,\quad \Theta\geq 2,\quad f_w(-y)=0.$$

Appendix 3 to Chapter III

Explicitly written expression (B 5) from the text is

$$(1)\qquad \psi\left(\frac{\mu}{\sigma}\right)\equiv\left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)}=-\frac{\alpha\left(\frac{\mu}{\sigma}\right)}{\beta\left(\frac{\mu}{\sigma}\right)},$$

where

$$\alpha\left(\frac{\mu}{\sigma}\right) \equiv \int_{-\infty}^{-\mu/\sigma} f_z(z) z U'\left(\frac{\mu}{\sigma}+z\right) dz + \int_{-\mu/\sigma}^{\infty} f_z(z) z U'\left(\frac{\mu}{\sigma}+z\right) dz, \tag{2}$$

$$\beta\left(\frac{\mu}{\sigma}\right) \equiv \int_{-\infty}^{-\mu/\sigma} f_z(z) U'\left(\frac{\mu}{\sigma}+z\right) dz + \int_{-\mu/\sigma}^{\infty} f_z(z) U'\left(\frac{\mu}{\sigma}+z\right) dz. \tag{3}$$

What is sought is the derivative $\psi'(\mu/\sigma)$. We first calculate

$$\begin{aligned}\alpha'\left(\frac{\mu}{\sigma}\right) = &\int_{-\infty}^{-\mu/\sigma} f_z(z) z U''\left(\frac{\mu}{\sigma}+z\right) dz - f_z\left(-\frac{\mu}{\sigma}\right)\left(-\frac{\mu}{\sigma}\right) U'(0-) \quad (=0) \\ &+ \int_{-\mu/\sigma}^{\infty} f_z(z) z U''\left(\frac{\mu}{\sigma}+z\right) dz + f_z\left(-\frac{\mu}{\sigma}\right)\left(-\frac{\mu}{\sigma}\right) U'(0+),\end{aligned} \tag{4}$$

$$\begin{aligned}\beta'\left(\frac{\mu}{\sigma}\right) = &\int_{-\infty}^{-\mu/\sigma} f_z(z) U''\left(\frac{\mu}{\sigma}+z\right) dz - f_z\left(-\frac{\mu}{\sigma}\right) U'(0-) \quad (=0) \\ &+ \int_{-\mu/\sigma}^{\infty} f_z(z) U''\left(\frac{\mu}{\sigma}+z\right) dz + f_z\left(-\frac{\mu}{\sigma}\right) U'(0+).\end{aligned} \tag{5}$$

Here $U'(0+)$ and $U'(0-)$ are the right-hand and the left-hand side derivatives of the gross-wealth utility function at a wealth level of zero. Note that

$$\psi'\left(\frac{\mu}{\sigma}\right) = -\frac{\alpha'\beta - \beta'\alpha}{\beta^2} \tag{6}$$

or, equivalently,

$$\psi'\left(\frac{\mu}{\sigma}\right) = \frac{\beta'}{\beta}\left(-\frac{\alpha'}{\beta'} + \frac{\alpha}{\beta}\right). \tag{7}$$

To calculate the size of this expression, consider first the quotient $-\alpha'/\beta'$. From the information on the first and second derivatives of the function $U(.)$ as given by expression (III B 2) in the text, we find

$$-\frac{\alpha'}{\beta'} = \frac{g+h}{i+j} = \frac{g}{i}\,\frac{i}{i+j} + \frac{h}{j}\,\frac{j}{i+j} \tag{8}$$

with

$$g \equiv -\int_{-\mu/\sigma}^{\infty} f_z(z) z \varepsilon(1-\varepsilon)\left(\frac{\mu}{\sigma}+z\right)^{-(1+\varepsilon)} dz, \tag{9}$$

(10) $$i \equiv \int_{-\mu/\sigma}^{\infty} f_z(z)\varepsilon(1-\varepsilon)\left(\frac{\mu}{\sigma}+z\right)^{-(1+\varepsilon)}dz,$$

(11) $$h \equiv -f_z\left(-\frac{\mu}{\sigma}\right)(1-\varepsilon)\frac{\mu}{\sigma}\lim_{\alpha\to 0+}\alpha^{-\varepsilon},$$

(12) $$j \equiv -f_z\left(-\frac{\mu}{\sigma}\right)(1-\varepsilon)\lim_{\alpha\to 0+}\alpha^{-\varepsilon}.$$

Here use has been made of

$$U'(0+) = \lim_{\alpha\to 0+} U'(\alpha) = (1-\varepsilon)\lim_{\alpha\to 0+}\alpha^{-\varepsilon}.$$

Consider now the quotient g/i. Its formal structure is the same as that of expression A calculated in appendix 2. (Substitute $g/i \equiv A$, $z \equiv w$, $\mu/\sigma \equiv y$, $f_z(.) \equiv f_w(.)$, $(1+\varepsilon) \equiv \Theta$ where $0<\varepsilon<1$, and reduce g/i by $\varepsilon(1-\varepsilon)$.) Hence, from expressions (24) and (28) in that appendix we have

$$\frac{g}{i}\begin{Bmatrix} = \\ < \end{Bmatrix}\frac{\mu}{\sigma} \quad \text{if} \quad f_z\left(-\frac{\mu}{\sigma}\right)\begin{Bmatrix} > \\ = \end{Bmatrix}0.$$

Further information on g/i is available from appendix 1. Expanding the quotient on the right-hand side of equation (4) in appendix 1 by $\varepsilon(1-\varepsilon)/x^{\varepsilon+\Delta}$, we find that $g/i=\psi_\Delta$ when $\Delta=1$. Now, from equation (III B 5) and equation (2) in appendix 1, we have $\psi = d\mu/d\sigma|_U$. In connection with condition ii) and statement (6) in appendix 1 this implies

$$\frac{g}{i} > \left.\frac{d\mu}{d\sigma}\right|_{U(\mu,\sigma)} \quad \text{if} \quad f_z\left(-\frac{\mu}{\sigma}\right) = 0.$$

With regard to expression (III B 7), we can therefore conclude

(13) $$\frac{g}{i} = \Gamma\left.\frac{d\mu}{d\sigma}\right|_U + (1-\Gamma)\frac{\mu}{\sigma},$$
$$1>\Gamma>0, \quad \text{if} \quad f_z(-\mu/\sigma)=0,$$
$$\Gamma=0, \quad \text{if} \quad f_z(-\mu/\sigma)>0.$$

Next consider the quotient h/j. Here we have the simple result

(14) $$\frac{h}{j} = \lim_{\alpha\to 0+}\frac{-f_z\left(-\frac{\mu}{\sigma}\right)(1-\varepsilon)\frac{\mu}{\sigma}\alpha^{-\varepsilon}}{-f_z\left(-\frac{\mu}{\sigma}\right)(1-\varepsilon)\alpha^{-\varepsilon}} = \frac{\mu}{\sigma}.$$

Taking account of $d\mu/d\sigma|_U = -\alpha/\beta$ from (1), noting that (5), (10), and (12) imply $\beta' = -(i+j)$, and utilizing (13) and (14), we can now write (7) in the form

$$(15) \qquad \psi'\left(\frac{\mu}{\sigma}\right) = -\frac{i+j}{\beta}\left\{\left[\Gamma\frac{d\mu}{d\sigma}\bigg|_U + (1-\Gamma)\frac{\mu}{\sigma}\right]\frac{i}{i+j} + \frac{\mu}{\sigma}\frac{j}{i+j} - \frac{d\mu}{d\sigma}\bigg|_U\right\}.$$

After some basic manipulation, this becomes

$$(16) \qquad \psi'\left(\frac{\mu}{\sigma}\right) = -\frac{i(1-\Gamma)+j}{\beta}\left(\frac{\mu}{\sigma} - \frac{d\mu}{d\sigma}\bigg|_U\right).$$

Now recall that in (12) the term $(1-\varepsilon)\lim_{\alpha\to 0+}\alpha^{-\varepsilon}$ stands for $U'(0+)$. Therefore, utilizing

$$(17) \qquad U'(0+) = -\int_0^\infty U''(z)\,dz = -\int_{-\mu/\sigma}^\infty U''\left(\frac{\mu}{\sigma}+z\right)dz$$

$$= \int_{-\mu/\sigma}^\infty \varepsilon(1-\varepsilon)\left(\frac{\mu}{\sigma}+z\right)^{-(1+\varepsilon)}dz$$

as well as (10), (12), and (13), we can transform equation (16) into

$$(18) \qquad \psi'\left(\frac{\mu}{\sigma}\right) = \frac{\varepsilon(1-\varepsilon)}{\beta}\int_{-\mu/\sigma}^\infty \left[f_z\left(-\frac{\mu}{\sigma}\right) - (1-\Gamma)f_z(z)\right]$$

$$\left(\frac{\mu}{\sigma}+z\right)^{-(1+\varepsilon)}dz\left(\frac{\mu}{\sigma} - \frac{d\mu}{d\sigma}\bigg|_U\right),$$

where

$$1 > \Gamma > 0, \quad \text{if} \quad f_z\left(-\frac{\mu}{\sigma}\right) = 0,$$

$$\Gamma = 0, \quad \text{if} \quad f_z\left(-\frac{\mu}{\sigma}\right) > 0.$$

Finally, we need to know which are the boundaries of β. These are given by

$$(19) \qquad 0 < \beta < \infty.$$

It is clear that β is stictly positive. The finiteness of β follows from appendix 2, since β has the same algebraic form as the denominator N

of expression (1) that was calculated with formulas (14)–(20). (Substitute $z \equiv w$, $\mu/\sigma \equiv y$, $U'(z+\mu/\sigma)=0$ for $z<-\mu/\sigma$, $U'(z+\mu/\sigma)=(1-\varepsilon)(z+\mu/\sigma)^{-\varepsilon}$ for $z>-\mu/\sigma$, $\varepsilon \equiv \Theta$.) The formulas refer to the case $f_z(-\mu/\sigma)>0$. Of course, β is, *a fortiori*, finite if $f_z(-\mu/\sigma)=0$.

Appendix 4 to Chapter III

The task is to check whether the integral

$$A=\int_{-y}^{\infty} f_w(y+w)(y+w)^{-\Theta}\,dw=\int_{0}^{\infty} f_w(w)w^{-\Theta}\,dw \tag{1}$$

is finite, where $\Theta>0$ and $f_w(w)$ is a function with the properties

$$0<|f_w'(0)|<\infty, \tag{2}$$

$$f_w(0)=0. \tag{3}$$

For the time being it is assumed that

$$\left|\int_0^{\infty} f_w(w)\,dw\right|<\infty. \tag{4}$$

An alternative assumption will be introduced in (15). Assume, moreover, that there is a number Δ, $\Delta>0$, such that

$$|f_w(w)|<\infty, \quad \text{if} \quad 0<w\le\Delta; \tag{5}$$

and split the integral (1) in the following way:

$$A=A_1+A_2, \tag{6}$$

$$A_1\equiv\int_0^{\Delta} f_w(w)w^{-\Theta}\,dw, \tag{7}$$

$$A_2\equiv\int_{\Delta}^{\infty} f_w(w)w^{-\Theta}\,dw. \tag{8}$$

Since, because of (4) and $\Theta>0$, it is obvious that $A_2<\infty$, we only have to consider A_1. Define rays throught the origin $\bar{b}w$ and $\underline{b}w$, $\bar{b}\ge\underline{b}>0$ such that

$$\underline{b}w\le|f_w(w)|\le\bar{b}w, \quad \text{if} \quad 0\le w\le\Delta. \tag{9}$$

Then we find the following boundaries for $|A_1|$:

(10) $$\lim_{\alpha\to 0+}\left\{\underline{b}\int_{\alpha}^{\Delta} w^{1-\Theta}\,dw\right\} \leq |A_1| \leq \lim_{\alpha\to 0+}\left\{\bar{b}\int_{\alpha}^{\Delta} w^{1-\Theta}\,dw\right\}.$$

By integration this yields

(11) $$\lim_{\alpha\to 0+}\left\{\frac{\underline{b}}{2-\Theta}(\Delta^{2-\Theta}-\alpha^{2-\Theta})\right\} \leq |A_1|$$

$$\leq \lim_{\alpha\to 0+}\left\{\frac{\bar{b}}{2-\Theta}(\Delta^{2-\Theta}-\alpha^{2-\Theta})\right\}, \quad \Theta \neq 2,$$

(12) $$\lim_{\alpha\to 0+}\{\underline{b}(\ln\Delta-\ln\alpha)\} \leq |A_1|$$

$$\leq \lim_{\alpha\to 0+}\{\bar{b}(\ln\Delta-\ln\alpha)\}, \quad \Theta = 2,$$

which implies

(13) $$\infty > \frac{\underline{b}}{2-\Theta}\Delta^{2-\Theta} \leq |A_1| \leq \frac{\bar{b}}{2-\Theta}\Delta^{2-\Theta} < \infty, \quad \Theta < 2,$$

(14) $$\infty \leq |A_1| \leq \infty, \quad \text{i.e.,}\ |A_1| = \infty, \quad \Theta \geq 2.$$

Consider now, instead of (4), the alternative assumption[1]

(15a) $$\left|\int_0^c f_w(w)\,dw\right| < \infty, \quad 0 < c < \infty,$$

(15b) $$|f_w(w)| < x < \infty, \quad w \geq c.$$

Since this assumption is weaker than (4), the condition $\Theta < 2$ continues to be required for finiteness. However, it only guarantees that $|\int_0^c f_w(w)\,w^{-\Theta}\,dw| < \infty$. The question is which values of Θ ensure that

(16) $$\left|\int_c^{\infty} f_w(w)\,w^{-\Theta}\,dw\right| < \infty.$$

[1] Note that, for the integral in equation (III B 9), assumption (15) is ensured, but not necessarily assumption (4).

Because of

$$(17)\qquad \left|\int_c^\infty f_w(w)w^{-\Theta}\,dw\right| < \left|x\int_c^\infty w^{-\Theta}\,dw\right|$$

$$= \lim_{w^*\to\infty}\left|x\int_c^{w^*} w^{-\Theta}\,dw\right|$$

$$= \lim_{w^*\to\infty}\begin{cases}|x(1-\Theta)(w^{*1-\Theta}-c^{1-\Theta})| \begin{cases}=\infty, & \Theta<1,\\ <\infty, & \Theta>1,\end{cases}\\ |x(\ln w^*-\ln c)| \qquad\qquad =\infty, \quad \Theta=1,\end{cases}$$

the answer is $\Theta > 1$. Thus the range of values of Θ that even under the weak assumption (15) implies a finiteness of the integral (1) is

$$(18)\qquad 1 < \Theta < 2.$$